

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2004, 13:18:42 ; Search time 5232 Seconds
(without alignments)
2925.208 Million cell updates/sec

Title: US-10-041-030-4

Perfect score: 2290

Sequence: 1 MRSFGGEHRCAPKPEVKYK.....ATOLVGEONCKILFQGPID 420

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O/cgnt 1/USPTO.spool.p/US10041030/runat.09122004.095517.7675/app.query.fasta.1.583
-DB=EST -OPMT=fastcap -SUFFIX=rest -MINMATCH=0.1 -LOOPEC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10041030.CGNT.1.1.6425.@runat.09122004.095517.7675 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=5
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hlc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|------------|--------------------|
| 1 | 2151.5 | 94.0 | 3387 | 3 AK030564 | AK030564 Mus muscu |
| 2 | 2147 | 93.8 | 1186 | 9 AY409116 | AY409116 Homo sapi |
| 3 | 2052.5 | 89.6 | 1183 | 9 AY409118 | AY409118 Mus muscu |
| 4 | 1926 | 84.1 | 3530 | 3 AK045515 | AK045515 Mus muscu |
| 5 | 1912 | 83.5 | 2584 | 3 AK045673 | AK045673 Mus muscu |
| 6 | 1681 | 73.4 | 1186 | 9 AY409117 | AY409117 Pan trogl |
| 7 | 1632 | 71.3 | 1879 | 3 AK044418 | AK044418 Mus muscu |
| 8 | 1626.5 | 71.0 | 1242 | 9 AY408790 | AY408790 Pan trogl |
| 9 | 1622 | 70.8 | 2572 | 3 CR606152 | CR606152 full-leng |

| | | | | | |
|----|--------|------|------|------------|--------------------|
| 10 | 1621.5 | 70.8 | 1242 | 9 AY408789 | AY408789 Homo sapi |
| 11 | 1611.5 | 70.4 | 1242 | 9 AY408791 | AY408791 Mus muscu |
| 12 | 1563 | 68.3 | 2474 | 3 AK033815 | AK033815 Mus muscu |
| 13 | 1519 | 66.3 | 849 | 5 BX388547 | BX388547 BX388547 |
| 14 | 1473 | 64.3 | 3531 | 3 AK029586 | AK029586 Mus muscu |
| 15 | 1430.5 | 62.5 | 957 | 7 CN329888 | CN329888 AGENCOURT |
| 16 | 1391 | 60.7 | 869 | 5 BX688136 | BX688136 BX688136 |
| 17 | 1363.5 | 59.5 | 856 | 6 CF241089 | CF241089 AGENCOURT |
| 18 | 1362 | 59.5 | 1117 | 7 CK027873 | CK027873 AGENCOURT |
| 19 | 1334.5 | 58.3 | 1076 | 4 BM800644 | BM800644 AGENCOURT |
| 20 | 1314 | 57.4 | 809 | 5 BX700555 | BX700555 BX700555 |
| 21 | 1306.5 | 57.4 | 744 | 7 BP682660 | BP682660 BP682660 |
| 22 | 1299 | 56.5 | 701 | 7 CK637810 | CK637810 UT-M-HO- |
| 23 | 1263.5 | 55.2 | 738 | 8 BP701801 | BP701801 BP701801 |
| 24 | 1261 | 55.1 | 768 | 7 CO738704 | CO738704 SILEBAC21 |
| 25 | 1258 | 54.9 | 784 | 5 BX749714 | BX749714 BX749714 |
| 26 | 1237 | 54.0 | 1229 | 3 BC026695 | BC026695 Mus muscu |
| 27 | 1203.5 | 52.6 | 752 | 5 BP709958 | BP709958 BP709958 |
| 28 | 1189 | 51.9 | 742 | 5 BX774759 | BX774759 BX774759 |
| 29 | 1185 | 51.7 | 1006 | 5 BQ049269 | BQ049269 AGENCOURT |
| 30 | 1161 | 50.7 | 766 | 7 CF539872 | CF539872 UT-M-BX- |
| 31 | 1159 | 50.6 | 889 | 5 BX327874 | BX327874 BX327874 |
| 32 | 1154.5 | 50.4 | 795 | 7 CK638809 | CK638809 UT-M-HO- |
| 33 | 1133 | 49.5 | 964 | 6 CD301176 | CD301176 AGENCOURT |
| 34 | 1126 | 49.2 | 826 | 4 BI183297 | BI183297 UNL-P-FN- |
| 35 | 1122 | 49.0 | 2342 | 3 AK013388 | AK013388 Mus muscu |
| 36 | 1106 | 48.3 | 949 | 5 BX369994 | BX369994 BX369994 |
| 37 | 1098 | 47.9 | 625 | 1 AL888857 | AL888857 AL888857 |
| 38 | 1089 | 47.6 | 893 | 5 BX410297 | BX410297 BX410297 |
| 39 | 1087 | 47.5 | 839 | 5 BU934702 | BU934702 AGENCOURT |
| 40 | 1083.5 | 47.3 | 659 | 1 AL873282 | AL873282 AL873282 |
| 41 | 1067.5 | 46.6 | 651 | 1 AL658565 | AL658565 AL658565 |
| 42 | 1065 | 46.5 | 876 | 6 CF251704 | CF251704 hdm003-CO |
| 43 | 1063 | 46.4 | 962 | 5 BX341733 | BX341733 BX341733 |
| 44 | 1061 | 46.3 | 593 | 4 BJ034307 | BJ034307 BJ034307 |
| 45 | 1052 | 45.9 | 600 | 4 BG803249 | BG803249 0222-86 M |

ALIGNMENTS

RESULT 1
AK030564
LOCUS
DEFINITION
AK030564 3387 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male pituitary gland cDNA, RIKEN full-length
sequence.
ACCESSION
AK030564
VERSION
AK030564.1 GI:26326560
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL
MEDLINE
99279253
PUBMED
10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
MEDLINE
20499374
PUBMED
11042155
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,

TITLE
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
Riken integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
20530913
11076861

REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

TITLE
JOURNAL
AUTHORS
5
The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
AUTHORS
6 (bases 1 to 3387)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsumi, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submision
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

TITLE
JOURNAL
COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://panom.gsc.riken.jp/
Location/Qualifiers
1. .3387
/organism="Mus musculus"
/mol_type="RNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:5330432G13"
/db_xref="taxon:10090"
/clone="5330432G13"
/sex="male"
/tissue_type="pituitary_gland"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
336. .1595
/note="unnamed protein product; peellino 2
[MGD|MGI:1891445, GB|NM_033602, evidence: BLASTN, 99%, match=864]
putative"
/codon_start=1
/protein_id="BAC27024.1"
/db_xref="GI:26326561"
/translation="MFSPGKKEPTPNKKPVKRELVVLYGYNGALPNGDRGRKSRFA
LYKRTYASGVSTIHMTVPQASKAISRGHSISYLTSLRQSTVAVETTHKDDMF
OYGRSTESPIDPVVDVTVSGGNEDAQIQTOSTISRPACILVDRNRPYARARIPAGFD
SKNRTFLGRKAAKWKPKPDGMDGLTTNGLVNHPQGGFLPESQPGWREISVCSGVYT
LAETRSAGQKGLVESETVLDDGSLIDCGATILMRWDGLFHAPTQKHIALQOEI
NBARQCPVGLNTLAFPSINRKEVVEEKPMAYLVSCGHVGHYSGHWSRSDTEANREBC

FEATURES
source

CDS

ORIGIN
PMCRITGVGFVPLMLGCEAGFYVDAGPPTHAFTPCGHVCSKSAKYTSQIPLPHGTAA
HAACFPACNTQLVGEQNKILFGQVVD"

Alignment Scores:
Pred. No.: 1,536-205 Length: 3387
Score: 2151.50 Matches: 395
Percent Similarity: 96.67% Conservative: 11
Best Local Similarity: 94.05% Mismatches: 13
Query Match: 93.95% Indels: 1
DB: 3 Gaps: 1

US-10-041-030-4 (1-420) x AK030564 (1-3387)

QY 1 MetPheSerProGlyGlnGlnGlnHisCySa1aProAnlyGluProValIleTyrgly 20
||||| :|||
Db 336 AGTTTTCCTCCCGGCANNAAGAACCCCAACAGANCCGGTGAATACCGG 395
|||||
QY 21 GluLeuValIleGlyTyraGnglyValaLeuProAnGlyAaspArglyYargArglys 40
|||||
Db 396 GAGCTGGTGTCTCTGGGTACATGTGCTTACTTAATGTGACAGGGGACAGGAAA 455
|||||
QY 41 SerArgPheAlaLeuTyrglyArgProlySa1aAnGlyValIleYpProSerThraHis 60
456 AGCAGATTGGCCCTATTAAGGAGCTTAAGCAGTGTGTCAACCCAGCACAAATCCAC 515
|||||
QY 61 ValIleSerThraProGlnAlaSerTyra1aIleSerCylyArgGlyGlnHisSerIleSer 80
516 ATGGCTCCACACCAAGCCGTCGAAGCCATCACTCCAGAGGACATACAGCATATCG 575
|||||
QY 81 TyrThrIleSerArgArgAnGlnThraValIleValIleGlyTyrThrHisAaspIleAaspThraP 100
576 TACAGCTGTGTACGAGGACCAAGCCGTAGTGTGAGTGCACACAGATTAAGACACAGC 635
|||||
QY 101 MetPheGlnValGlyArgSerThraGluSerProIleAaspPheValIleThraPThrIle 120
636 ATGTTTCAGTGGTGGCAGGTCAACAGAAAGCCCATGTGCTGTGTGCACAGACACGGTT 695
|||||
QY 121 SerGlySerGlnAnThraPaspGluIleGlnIleThraGlnSerThraIleSerArgPheAla 140
696 TCCGCGGTCAAGAACAAAGAT--GCCAGATCACACAGAGCAATCTTAGTTCGGA 752
|||||
QY 141 CysAlaGlyLevalCyAaspArgAnGluProTyrrAlaArgIlePheAlaIleAglPhe 160
753 TCCAGAGATCGTGTGACAGAGAACGACCATACACAGACGATTTTGGCGGCAAGATTCC 812
|||||
QY 161 AaspSerTyraPheAnIlePheIleGlyGlySa1aIleAlaIleTyrrIleAaspProAaspGly 180
813 GATTCCTCCAAAATAATCTTCTTGTGAGAGAAAGCAGAAATGAGAAAACCTCATGGA 872
|||||
QY 181 HisMetAaspGlyLeuThraPaspGlyValIleuValIleMetHisAProArgIlyGlyPheThr 200
873 CACATGAGATGACCTACTACCAACGGTGTCTGGAGTAGCACCCGACAGAGGCTTCACC 932
|||||
QY 201 GluGluSerGlnProGlyValIleThraPaspGlyIleSerValCySgIyAaspValIleTyrrIleu 220
933 GAGGATCCCAAGCCCTGAGGTCTGAGAGAGATCTGTCTGTGGGAGTATATACACCTTG 992
|||||
QY 221 ArgGlyThrArgSerAlaGlnGlnArgGlyIleuValIleGluSerGlnThraPheValIleu 240
993 CGAGAGACACAGTCCGGCCAGACAGAGGAGAAAGCTGTGTGAGAAAGTGAACCAACTCTCGT 1052
|||||
QY 241 GlnAaspGlySerIleuIleAaspLeuCySgIyIleThraIleuThraPaspGly 260
1053 CAAGACGGCTCCCTATGACCTGTGTGGGGCACTCTCTGTGAGAAACGCAATGAC 1112
|||||
QY 261 LeuPheHisThrProThraGlnIleHisIleGlnIleAlaLeuAArgGlnGlnIleAasnIleAla 280
1113 CTTTTCACGCTCCCTACTACAGAGACATTAAGGCTCCGGCAGAGATGATATCAGCC 1172
|||||
QY 281 ArgProGlnCyAProValGlyIleuAnThraIleuAlaPheProSerIleAasnArgIlySgIy 300
1173 GCACCCCAAGTCCCGTGGGCTTAAACCTGTGCTTCCAGCATCAACCGAGAGGA 1232
|||||
Db

| | | | |
|------------------------|--|--|------------------------------|
| Qy | 301 | ValValgluGlulysValGlnProTTPalaLeTyrLeuSerCysValGlnValAlHIGlTYTThIS | 320 |
| Db | 1223 | GTGTGTGAGAGAGAGACACCCTTGCGCATTACTGAGCTGGGCATGTGTACAGGCTTACAC | 12322 |
| Qy | 321 | AenTPrgjYhIaTgSersAspThrGlnIuaIaaenGuaTgIuCySPrometCysaTgThr | 340 |
| Db | 1293 | AGCTGGGGGCATCCGAGCGGACACCGGAGCCAAACGAGAGGAGTGTCCTATGTGACGACT | 13522 |
| Qy | 341 | ValGlyProTyrValProLeuTrrleuGlyCysGlnIaGlyPheTyrValAspAlaGly | 360 |
| Db | 1353 | GTGGGCCCCCTACGTCCTCTCTGCTGGCTGTGTGAGCAGATTTTATGTGATCGCGGA | 1412 |
| Qy | 361 | ProProThrHIsAlaPheThrProCysGlyHIsValCysSerGlnLysSerAlaLysTyr | 380 |
| Db | 1413 | CCCCCACTACGCGCTTTACACCCCTCGGGGACGCTGTTCAGAGAACTTTCGCAAGTAC | 14722 |
| Qy | 381 | TrrSerGlnIleProLeuProHIsGlyThrrHIsAlaPheHIsAlaAlaCysProPheCys | 400 |
| Db | 1473 | TGTCGTGAGATCCGACTGCGCCCCACGAGAGCAGCCGGTTTCAAGCGCGCTGTCCGTTGCG | 15322 |
| Qy | 401 | AlaThrGlnLeuValGlyGlnIuaenCysIleLysLeuIlePheGlnGlyProIleAsp | 420 |
| Db | 1533 | GCCACGCGAGCTGGTGTGTGAACAGAACTGATCAATGATTTTCCAAAGTCCAGTGCAC | 15922 |
| RESULT 2 | | | |
| LOCUS | AY409116 | 1186 bp | DNA linear GSS 16-DEC-200303 |
| DEFINITION | Homo sapiens PEL12 gene, VIRUAL TRANSCRIPT, partial sequence, | | |
| ACCESSION | AY409116 | | |
| VERSION | AY409116.1 GI:39765084 | | |
| KEYWORDS | GSS. | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | |
| AUTHORS | 1 (bases 1 to 1186) Clark, A.G., Gianoweki, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tenenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Slnsky, J.J., Adams, M.D. and Cargill, M. | | |
| TITLE | Interfing nonneutral evolution from human-chimp mouse orthologous gene trics | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | |
| REFERENCE | 14671302 | | |
| AUTHORS | 2 (bases 1 to 1186) Clark, A.G., Gianoweki, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tenenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Slnsky, J.J., Adams, M.D. and Cargill, M. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | |
| COMMENT | This sequence was made by sequencing genomic exons and ordering them based on alignment. | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..1186 | | |
| gene | /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>1186 /gene="PEL12" /locus_tag="HCM3450" | | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Pred. No.: | 8.77e-206 | Length: | 1186 |
| Score: | 2147.00 | Matches: | 394 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 93.76% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-10-041-0930-4 (1-420) x AY409116 (1-1186)

| | | | |
|----|-----|---|------|
| QY | 27 | TyrAenGIyAlaLeuProhEnGIyAaPrgGIyArGaTgLySeSaArghPhaAlaLeuTyr | 46 |
| Db | 2 | TACATAGGTGCTTTACCCAAATGGAATGAGAGAGGAGGAAATAAGATTGGCCCTCAC | 61 |
| QY | 47 | LysArpProLySaIaAenGIyValLyProSeArThVaIhSaVAlIeSeArThProGln | 66 |
| Db | 62 | AAAGGAGCCCAAGGCAATGTGTCAAAACCAGACCGTCCATGTGATTCACAGCCCGAG | 121 |
| QY | 67 | AlaSeArLySaIaIeSeArCyAaGIyGlnhIaSeArIeSeArTyrThLeuSeArGln | 86 |
| Db | 122 | GCATTCAAAGCTTACAGCTGCAGAAAGTCAACACAGTATATCTTACACTTGTCAAGGAT | 181 |
| QY | 87 | GlnThVaIValValGIyTyrThhIaAPlLyAaPThrAaPMeCPhieGlnValGIyArG | 106 |
| Db | 182 | CAGACTGGTGGTGGAGGTACACACATGATAGATACGGATATGTTTCAGTGGCAGA | 241 |
| QY | 107 | SeArThGlnSeArProIleAaPPhVaIValThraPThrIleSeArGIySeArGlnAaThr | 126 |
| Db | 242 | TCAACAGAAACCCTTATGCACTTGTGTGCACACACAGATTTCTGGCAGCCAAACAG | 301 |
| QY | 127 | AaPGLuaGIaGlnIleThrGlnSeArThrIleSeArghPheaIaCyAaGIyValCyAaP | 146 |
| Db | 302 | GACGAAGCCAGATCACACAGAGACCATATCCAGTTCGGCTCAGGATGCTGTGGAC | 361 |
| QY | 147 | ArGaenGIyProTyrThrAlaArGIlePheaIaIaGIyPheaPSeArSeLySaAaIle | 166 |
| Db | 362 | AGGAATGAACCTTACACAGCAGGATATTTCGCCCGGATTGACTTCCAAAAACATA | 421 |
| QY | 167 | PhaenGIyGlnLySaIaAaLyTTrLyAaPProAaPGLyNhIaMeArPGLyLeuThr | 186 |
| Db | 422 | TTTCTGGAGAAAAGGCAAGATGGAAAAACCCAGCGCCACATGGATGGGCTCACT | 481 |
| QY | 187 | ThraenGIyValLeuValMeChIaProArGIyGIyPheThrGlnGlnSeArGlnProGIy | 206 |
| Db | 482 | ACTAATGAGCGTCTGTGTATCATCCACAGAGGGGCTTACAGAGAGTCCCAACCCCGAG | 541 |
| QY | 207 | ValTTrPAaPGLyIleSeArLyAaPValTyrThrIleuArGIyThThArGSeArIa | 226 |
| Db | 542 | GTCGCGCCGAGATCTGTCTGTGGAGATGTACACCTTGCAGAAACCAAGTCCGCC | 601 |
| QY | 227 | GlnGlnArGIyLyLyLeuValGIuSeArIuThraenValLeuGlnAaPGLySeArIle | 246 |
| Db | 602 | CAGCAACAGGAAGAACTGTGTGAAAGTGAACCAACGCTCCGACAGACGGCTCTCATTT | 661 |
| QY | 247 | AaPLeuCyGILyAlaThrLeuLeuTTrArGIhIaAaPGLyLeuPhehIeThrProThr | 266 |
| Db | 662 | GACCTGTGTGGGCCACTCTCTCTGGAGAACACAGATGGGCTTTTTCATATCCCAACT | 721 |
| QY | 267 | GlnLyhIaIleGIyAlaLeuArGIuGlnGlnIleAaPGLyAaArProGlnCyProVal | 286 |
| Db | 722 | CAGAAAGCAATAGAAGCCCTCCGGCAGAGATTAAACCGCCCGCCCTCAATATGCTGTG | 781 |
| QY | 287 | GIyLeuAaenThrLeuAaPheProSeArIleAaArGIyGIyValValGIuGlnLySeGln | 306 |
| Db | 782 | GGGCTCAACACCTCGGCTTCCCAAGCATCAACAGAAAGGTGTGTGGAGAAAGAG | 841 |
| QY | 307 | ProTTrPAaLyTTrLeuSeArCyGILyValhIaIleGIyTyrhIaAaenTTrPGLyNhIaArGSe | 326 |
| Db | 842 | CCCTGGGCAATATCTCACTTGTGGCCAGCTGCACGGGTACCAACTGGGGCCATCGAGT | 901 |
| QY | 327 | AaPThrGIyAlaAaenGIyArGIuCyProMeArCyAaArGIyThVaIIGlyProTyrValPro | 346 |
| Db | 902 | GACACGAGGCGCAACAGAGGAGGTGTCCATGTGCAGAACTGTGGGCCCTATATGTGCT | 961 |
| QY | 347 | LeuTTrPAaenGIyCyGILyAlaGIyPheTyrValAaPGLyProToThrNhIaAaPhe | 366 |
| Db | 962 | CTCTGGCTGTGGCTGTAGAGCAGGATTTATGTAAACGAGGACCGGCAACTCATGTCTTC | 1021 |
| QY | 367 | ThrProCyGILyNhIaValCySeArGIyLySeArAlaLyTyrTTrPSeArGlnIleProLeu | 386 |

| | | | |
|--|--|--|-----------------|
| Db | 1022 | ACTCCCTGTGACACAGCTGTGCTCGGAGAAAGTCTGCAAAATATCTGCTCAGATCCCGTTG | 1081 |
| Oy | 387 | ProHsiGlyYThrHisAlaPheHisAlaAlaCysProPheCysAlaThrGlnLeuValGly | 406 |
| Db | 1082 | CCTCATGTGAACATCAATGACATTCACCGCTCTTGCCCTTCTGTGTGCTACACAGCTGGTTGGG | 1141 |
| Oy | 407 | GLuGlnAnCyValIleValLeuIleIlePheGlnGlyProIleAsp | 420 |
| Db | 1142 | GAGCAAAACTGCATCAATTAATTAATTTCCAAAGGTCCAATTGAC | 1183 |
| RESULT 3 | | | |
| AY409118 | | 1183 bp | DNA |
| LOCUS | | | linear |
| DEFINITION | Mus musculus PELI2 gene, VIRUTUAL TRANSCRIPT, partial sequence, | | GSS 16-DEC-2003 |
| ACCESSION | AY409118 | | |
| VERSION | AY409118.1 | | GI:39765086 |
| KEYWORDS | GSS. | | |
| SOURCE | Mus musculus (house mouse) | | |
| ORGANISM | Mus musculus | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| AUTHORS | 1 (bases 1 to 1183) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | |
| TITLE | Interfering nonneutral evolution from human-chimp-mouse orthologous gene trios | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | |
| PUBMED | 14671302 | | |
| REFERENCE | 2 (bases 1 to 1183) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | |
| AUTHORS | Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | |
| JOURNAL | | | |
| COMMENT | This sequence was made by sequencing genomic exons and ordering them based on alignment. | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..1183 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..>1183 /gene="PELI2" /locus_tag="HGM3450" | | |
| gene | | | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Pred. No.: | 3,07e-196 | Length: | 1183 |
| Score: | 2052.50 | Matches: | 375 |
| Percent Similarity: | 97.72% | Conservative: | 10 |
| Best Local Similarity: | 95.18% | Mismatches: | 8 |
| Query Match: | 89.63% | Indels: | 1 |
| DB: | 9 | Gaps: | 1 |
| US-10-041-0350-4 (1-420) x AY409118 (1-1183) | | | |
| Oy | 27 | TyrAnGlyAlaLeuProAsnGlyYAspArgGlyArgTgLYSserArgPheAlaLeuTyr | 46 |
| Db | 2 | TACATAGTGTCTTACTCTAATGTGTACAGGGGACAGAGAAAGACAGATTGGCCCTCAT | 61 |
| Oy | 47 | LYsArGProLYsAlaAsnGlyValIleYAspRoserThrValHisValIleSerThrProGln | 66 |
| Db | 62 | AAGCGGACCTACGCGCAGTGTGTCAAAACCAACGACAAATCCATGCTCTCCACACACAG | 121 |
| Oy | 67 | AlAserLYsAlaIleSerCYsYsIleGlyGlnHisSerIleSerTYrThrLeuSerArgAsn | 86 |
| Db | 122 | GGGTTCAGAGGCGCATCAGTCAAGGACGACATCAAGCATATGTGTACAGTGTACGGAGC | 181 |

| LOCUS | DEFINITION | AK045515 | 3530 bp | mRNA | linear | HTC 03-APR-2004 |
|-------|------------|----------|---|------|--------|-----------------|
| QY | | 87 | GlnthrtValValValGluTyrTrpThrHisAspLysAspSerThrAspMetPheGlnValGlyArg | 106 | | |
| Db | | 182 | CAGACGGTAGTGGTGGAGTACACACAGATTAAGACAGGACATGTTCCAGTGGGACGG | 241 | | |
| QY | | 107 | SerThrGlnSerProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThr | 126 | | |
| Db | | 242 | TCAACAGAAAGGCCCATTTGACTTCGTGTGTACAGACACGGTTTCCGGCGGTGAGAAAGAA | 301 | | |
| QY | | 127 | AspGluValaglnIlethrglnSerThrIleSerArgPheAlaCysArgIleValCysAsp | 146 | | |
| Db | | 302 | GAT---GCCCGATCATCACAGAGACACCATCTCTGTGGTTCGATGCATGCATGCTGTGCAC | 358 | | |
| QY | | 147 | ArgAsnGluProTyrTrpThrAlaArgIlePheAlaIleGlyPheAspSerSerLysAsnIle | 166 | | |
| Db | | 359 | AGGACAGACGCATATACAGACCGCATATATGCGCGGACAGATTCGATTTCCAAATAATATC | 418 | | |
| QY | | 167 | PheLeuGlyGluLysAlaIleValTyrTrpLysAsnProAspGlyHisMetAspGlyLeuThr | 186 | | |
| Db | | 419 | TTTCTTGGAGGAAAGACAGCAAAATGAAAAACCTTATGACACATGATGACTACT | 478 | | |
| QY | | 187 | ThrAsnGlyValLeuValMetHisProArgGlyGlyPheThrGlnGlnSerGlnProGly | 206 | | |
| Db | | 479 | ACCAATGGTGTCTAGTATGATGACACCCCGAAGAGAGGCTTACACGAGAAATCCACAGCTGGA | 538 | | |
| QY | | 207 | ValTrpArgGluIleSerValCysGlyAspValTyrThrLeuAspGluTrpArgSerAla | 226 | | |
| Db | | 539 | GTCGTGAGAGAGATCTGTCTGTGGGAGATGTATACCTTTCGACAGACACAGGTCTGGCC | 598 | | |
| QY | | 227 | GlnGlnArgGlyLysLeuValGlnSerGlyThrAsnValLeuGlnAspGlySerLeuIle | 246 | | |
| Db | | 599 | CAGCAGAGGGGAAAGCTGTGTGGAAAGTGAAGCACAGCTCTGCAGACGGCTCCTTATT | 658 | | |
| QY | | 247 | AspLeuCysGlyAlaThrLeuLeuTrpArgThrAlaAspGlyLeuPheHisThrProThr | 266 | | |
| Db | | 659 | GACCTGTGTGGGGCCACTCTCTCTGTGAGAAACCGCAGATGGCTTTTTCACAGCTCTACT | 718 | | |
| QY | | 267 | GlnLysHisIleGluValAlaLeuArgGlnGluIleAsnAlaAlaArgProGlnCysProVal | 286 | | |
| Db | | 719 | CAGAAAGACATAGAGAGCCCTCCCGCAGAGATCATGACGCCACCCACAGTGGCCCGTG | 778 | | |
| QY | | 287 | GlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValValGlnLysGln | 306 | | |
| Db | | 779 | GGCTTTACACCTTGCCCTTCCCGACATCAACCGAAGAAAGTGTGGAAAGAAAGACAG | 838 | | |
| QY | | 307 | ProTrpAlaTyrLeuSerCysGlyHisValHisGlyTyrHisAsnTrpGlyHisArgSer | 326 | | |
| Db | | 839 | CCCTGGGCGATACCTGACACTGGGCGCATGTGACGGCTACACAGCTGGGGCGATCGGAGC | 898 | | |
| QY | | 327 | AspThrGluAlaAsnGluArgGluCysProMetCysArgThrValGlyProTyrValPro | 346 | | |
| Db | | 899 | GACCACGGAAGCAACGAGAGGAGTGTCCCATGTGCGAGACTGTGGGCCCTTACGTCTCT | 958 | | |
| QY | | 347 | LeuTrpLeuGlyCysGlyValaglyPheTyrValAspAlaGlyProProThrHisAlaPhe | 366 | | |
| Db | | 959 | CTGTGGCTGGCTGTGAGCGACGATTTTATGTTCATGCGGAGACCCCAACTCATACGCTTTC | 1018 | | |
| QY | | 367 | ThrProCysGlyHisValCysSerGlnLysSerAlaLysTyrTrpSerGlnIleProLeu | 386 | | |
| Db | | 1019 | ACCCCTGGCGGACACGCTCTTTCAGAAAGTCTGCCAAGTACTGTGTGCAGATCCCACTG | 1078 | | |
| QY | | 387 | ProHisGlyThrHisAlaPheHisAlaAlaCysProPheCysAlaThrGlnLeuValGly | 406 | | |
| Db | | 1079 | CCCCACGGAAGCAGCAGCGTTTCAAGCGGCTGTCCGTTTCGCGCAGCAGAGCTGTGGT | 1138 | | |
| QY | | 407 | GlnGlnAsnGlySerIleLysLeuIlePheGlnGlyProIleAsp | 420 | | |
| Db | | 1139 | GAAACGAATGCATCAATAATGATTTTCCAAAGTCTCACTGAGAC | 1180 | | |

QY 201 GUGUGUSeGlnProGlyValTTPArGluIleSeVaiCyseGlyAsePValTYrThrIleu 220
 Db 759 GAGGATTCACAGCTCGAGTCTGAGAGAGATCTCTGCTGGGAGTGTATACCTTG 818
 QY 221 ArgGluThrArgSerAaGlnGlnAArgGlyLysLeuValGluSeGlyLuthAenValIleu 240
 Db 819 CGAGAGACACAGGTCGGCCAGCAGAGAGGAGAAAGCTGTGGAAAGTAGACCAACGTCCTG 878
 QY 241 GlnAseGlySerIleuLeuLeuLeuCyseGlyValAThrLeuLeuThrArgThrAaAsePgly 260
 Db 879 CAGAGAGGCTCCCTCATATGACCTGTGTGGGCGACCTCTCTGAGAAATCGCAGATGGC 938
 QY 261 LeuPheHisThrProThrGlnLysHisGileGluAlaLeuAArgGlnGlnIleAenAlaIa 280
 Db 939 CTTTTCACGCTCTCTACTAGAGACATATGAAAGCCCTCCGAGAGAGATCAATGACGCC 998
 QY 281 ArgProGlnCyseProValGlyLeuAenThrLeuAlaPheProSerIleAenAArgLysGlu 300
 Db 999 CGAGCCACAGTGGCCGCTGGGCTTAAACCCCTGGCTTCCCGCAGCATCAACCGAAGAA 1058
 QY 301 ValValGluGluLysGlnProThrPalATyrLeuSerCyseGlyHisValHisGlyTyrHis 320
 Db 1059 GTGTGTGAAGAAGAGACGCTGGGAGTACCTGAGCTGGCCATGTGACAGGCTACAC 1118
 QY 321 AenThrGlyHisArgSerAsePThrGluAlaAenGluAArgGluCysePrometCyseArgThr 340
 Db 1119 AGCTGGGGCCATCGAGACACAGAGACAGAGAGAGAGTCTCCATGTGCAGAGACT 1178
 QY 341 ValGlyProTyrValProLeuThrLeuGlyCyseGluAlaGlyPheTyrValAaPalagly 360
 Db 1179 GTGGGCGCTACCTCTCTGTGGCTGTGAGAGAGAGATTTATGTGATGTGGGGA 1238
 QY 361 ProProThrHisAlaPheThrProCyseGlyHisValCyseSerGluLysSerAlaLysTyr 380
 Db 1239 CCCCACACTCAGCTTTCACCCCTCGGGCAGCTGTGTTCAGAGAAGTCTGCCAAGTAC 1298
 QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCyseProPheCys 400
 Db 1299 TGTGTCCAGATCCCACTGCCGCCAGAGACGACGCGCTTTCAGCCGCCCTGTCCGTCTGC 1358
 QY 401 AlaThrGlnLeuValGlyGlnAenCyseGlyLeuLeuLeuPheGlnGlnProIleAseP 420
 Db 1359 GCCACGACAGCTGTGTGTGAACAGACTGCATCAATTTTCCAAAGTCCAGTGCAC 1418
 RESULT 5
 LOCUS AK045673 2584 bp mRNA linear HTC 03-APR-2004
 DEFINITION Mus musculus adult male corpora quadrigemina cDNA, RIKEN
 full-length enriched library, clone:B230220E21 product:pellino 1,
 full insert sequence.
 ACCESSION AK045673
 VERSION AK045673.1 GI:26337532
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subjection of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 JOURNAL sequencing pipeline with 384 multiplexed sequencer
 MEDLINE Genome Res. 10 (11), 1757-1771 (2000)
 PUBMED 20530913
 REFERENCE 11076861
 AUTHORS 4
 TITLE The RIKEN Genome Exploration Research Group Phase II Team and the
 JOURNAL PANTOM Consortium.
 REFERENCE Functional annotation of a full-length mouse cDNA collection
 AUTHORS Nature 409, 685-690 (2001)
 TITLE 5
 JOURNAL The PANTOM Consortium and the RIKEN Genome Exploration Research
 REFERENCE Group Phase I & II Team.
 AUTHORS Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 REFERENCE 6 (bases 1 to 2584)
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanakawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akhiba, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp,
 URL: http://genome-gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 COMMENT cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome-gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/
 URL: http://fantom.qualifiers
 FEATURES
 source
 1. 2584
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM.DB:B230220E21"
 /db_xref="taxon:10090"
 /clone="B230220E21"
 /sex="male"
 /tissue_type="corpora quadrigemina"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 426. 1682
 /note="unnamed protein product; pellino 1
 (MGDI:1914495, GB|NM_023324, evidence: BLASTN, 99%,
 match=2228)
 putative"
 /codon_start=1
 /protein_id="BAC32452.1"
 /db_xref="GI:26337533"
 /translation="MFSPDQENHPKAPVKYGEILVLGNGSLPNGDRGRKRSFALP

| | | | |
|-------------------------|---------|-------------------|-----|
| Score: | 1693.00 | Matches: | 317 |
| Percent Similarity: | 80.46% | Conservative: | 0 |
| Best Local Similarity: | 80.46% | Mismatches: | 77 |
| Query Match: | 73.41% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |
| US-10-041-030-4 (1-420) | x | AY409117 (1-1186) | |

| | | | |
|----|-----|---|-----|
| QY | 2 | TYRASNGLVALALEUPROANGIYASPARAGGIYATAGATGYLSESSERARGPHEALALEUTYR | 46 |
| DB | 2 | THCAATGTCCTTTACCCTCAATGAGATGAAGACGAGGAAAAGTAGATTTGGCCCTCTAC | 61 |
| QY | 47 | LYSARGPROLYSALAANGIYVALLYSPROSETRVALHISEVALISESTRPROGLN | 66 |
| DB | 62 | AAGGGGCCCAAGGCAATGGTGTCAAAACCACGACTGTCCATGTGATATCCACGCCCCAG | 121 |
| QY | 67 | ALASERLYSALAILESERCYSLYSGIYGLNHISERLISESTRYTHRLEUSERARGSN | 86 |
| DB | 122 | GCGTCCAAAGCTTCATGACTGTCAAAGGTCAAACAGATATCTCTACACTTGTCAAGGAAT | 181 |
| QY | 87 | GLNTHRALVALAIGUITYRTHHISABRYLSAPRTHASPMERPHGLNVALGLYARG | 106 |
| DB | 182 | CAGCTGTGGTGGGATGACACATGATGAAGACACGGAATATTTCCAGNNNNNNNN | 241 |
| QY | 107 | SETRHIGUSERPROILLEASPHEVALITHRAPRTHILESERGLYSERGLNASTHR | 126 |
| DB | 242 | NN | 301 |
| QY | 127 | ASPGIUALAGLNILETHRGINSETRHILESERARGPHEALCYAARGILEVALCYAEP | 146 |
| DB | 302 | NN | 361 |
| QY | 147 | ARGASGLUPROTyrTHRHALARGILEPHEALAGIYpHEASPSERLYASAHLILE | 166 |
| DB | 362 | NN | 421 |
| QY | 167 | PHLEUGLYLULYSALAALALYSTRLYSAPPROASPGLYHISEMETSASPGLYLEUTHR | 186 |
| DB | 422 | NNNNNNNNNNAAAAAGGACGAAAGTGGAAAAACCAGCGGCACATAGATGGGCTCACT | 481 |
| QY | 187 | THRASNGLYVALLEUVALMECHISPROARGIYGLYPHETHRGLUGUSERGLNPROGLY | 206 |
| DB | 482 | ACTAATGGCCTTCGTGGATGATCCACACGAGGGGCTTCCACGAGAGTCCCAACCCGGG | 541 |
| QY | 207 | VALTRPRGGLULESERVALCYSGIYASBPVALTYRTHLEUARGGLNTHRARGSERALA | 222 |
| DB | 542 | GTCGTGGGCGAGATCTGTCTGTGAGATGTGTACACTTGCAGAAACAGAGTGGGCC | 601 |
| QY | 227 | GLNGLNARGLYLYSLEUVALIGUSERGLUTHRAENVALLEUGLNASPGLYSERLEUILE | 246 |
| DB | 602 | CAGCAACGAGGAAAGCTGGTGGAAATGAGACCAAGCTCTGCAGAGAGGGCTCCTCATTT | 661 |
| QY | 247 | ASPLEUCYSGLYALATHRLEULEUTPRATGTHRALASBPGLYLEUPHEHISTHRPROTHR | 266 |
| DB | 662 | GACCTGTGTGGGGCACTCTCCCTCGAAMACAGACAGATGGGCTTTTCCATATCCCAACT | 721 |
| QY | 267 | GLNLYSISIRLEGLUALALEUARGINGIYULEANALALARGPROGLNICYSPROVAL | 286 |
| DB | 722 | CAGAGACATPAGAGGCTTCGCGGACGAGATTAAGCCGCCCGGCTCAGATGTCGTG | 781 |
| QY | 287 | GLYLEUANRTHRLEUALAPHEPROSERILEANARGLYSGIULVALIGIUGIULYSGLN | 306 |
| DB | 782 | GGGCTCAACACCCGTGGCTTCCCAACATCAACAGGAAGAGTGTGTGAGAGAGAACAG | 841 |
| QY | 307 | PROTRPALATYRLEUSERCYSGIYHISEVALHISGLITYRHHISAEHTPRGLYHISARGSER | 326 |
| DB | 842 | CCCTGGGCAATATCTCAGTGTGTGGCCACGTGCACGGGTATCCCAACTGGGGCCATCGAGT | 901 |
| QY | 327 | ASPRTHGLUALAANGIULARGIYCYSPROMETCYAARGTHRVALGLYPROTYRVALPRO | 346 |
| DB | 902 | GACACGGAGGCCAACGAGAGGAGTGTCCNNNNNNNNNNNNNNNNNNNNNNNNNNNN | 961 |
| QY | 347 | LEUTRPLEUGLYCYSGIUALAGIYPHETYRVALASPALAGIYPROTOTHRHISALAPHE | 366 |

| | | | | |
|------------|---|---------|---|------------------------|
| D6 | | 962 | CTCAGGCTGGCTGTGTGAGGAGATTTTATTATACCGACGACCAGCCAACTNNNGCTTC | 1021 |
| OY | | 367 | ThPrProcysglYhiSvAlCySeSerGluySeSerAlayStYrTrPseRginIlleProLeu | 386 |
| D6 | | 1022 | ACTCCCTGTGACACGCTGTCTCGAGANAAGTCTGCAAAATACGTGTTCAGATCCGCTTG | 1081 |
| OY | | 387 | PRHISgLyThrHisAlaPhenHiSaIeAlaAlaCYSeProBheCYbaLalThnGIInLeuValIGly | 406 |
| D6 | | 1082 | CCTCATGGAACtCAcTGAcATTTCACGCTGTCGCCCTTCTGTGTCTACACAGCTGGTGGG | 1141 |
| OY | | 407 | GLInLanPmCYstIelySeIuIllePhghGngLYProIleAap | 420 |
| D6 | | 1142 | GAGCAAAACtGCATCAATAATTATTTTCCAAGTTCCAATTGAC | 1183 |
| RESULT 7 | | | | |
| LOCUS | AK044418 | 1879 bp | mRNA | linear HTC-03-APR-2004 |
| DEFINITION | Mus musculus adult retina cDNA, RIKEN full-length cDNA cloning project; similar to PELLINO 2 (PELLINO [DROSOPHILA HOMOLOG 2] [Homo sapiens], full insert sequence). | | | |
| ACCESSION | AK044418 | 1 | GI:26336455 | |
| VERSION | AK044418.1 | | | |
| KEYWORDS | HTC; CAP trapper. | | | |
| SOURCE | Mus musculus (house mouse) | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| REFERENCE | | 1 | Carninci, P. and Hayashizaki, Y. | |
| AUTHORS | | | | |
| TITLE | High-efficiency full-length cDNA cloning | | | |
| JOURNAL | Meth. Enzymol. 303, 19-44 (1999) | | | |
| MEDLINE | 99279253 | | | |
| PUBMED | 10349636 | | | |
| REFERENCE | | 2 | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komio, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. | |
| AUTHORS | | | | |
| TITLE | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes | | | |
| JOURNAL | Genome Res. 10 (10), 1617-1630 (2000) | | | |
| MEDLINE | 20499374 | | | |
| PUBMED | 11042159 | | | |
| REFERENCE | | 3 | Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Komio, H., Akiyama, J., Niishi, K., Kitamura, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Mashino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasaiyagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunuma, S., Kawai, U., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y. | |
| TITLE | RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer | | | |
| JOURNAL | Genome Res. 10 (11), 1757-1771 (2000) | | | |
| MEDLINE | 20530913 | | | |
| PUBMED | 11076661 | | | |
| REFERENCE | | 4 | The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. | |
| AUTHORS | | | | |
| TITLE | Functional annotation of a full-length mouse cDNA collection | | | |
| JOURNAL | Nature 409, 685-690 (2001) | | | |
| REFERENCE | | 5 | The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. | |
| AUTHORS | | | | |
| TITLE | Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs | | | |
| JOURNAL | Nature 420, 563-573 (2002) | | | |
| REFERENCE | | 6 | (bases 1 to 1879) | |
| AUTHORS | | | | |
| TITLE | Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Hayashida, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiwamoto, K., Hiraoaka, T., Hirokane, T., Hoti, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Katoh, H., Kawai, U., Kojima, Y., Kondo, S., Komio, H., Kouda, M., Koya, S., Kurihara, C., Matsumura, T., Miyazaki, A., Murata, M., | | | |

TITLE
JOURNAL

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Kahira, S., Takekida, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

COMMENT

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216).

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Retina RNA was provided by Dr. Stefano Gualtinchich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES

source

1. 1879

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:A930011L17"

/db_xref="taxon:10090"

/clone="A930011L17"

/tissue_type="retina"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

104..1441

/note="unnamed protein product; putative

similar to PELLINO 2 (PELLINO 2 (DROSOPHILA) HOMOLOG 2)

[Homo sapiens] (SPTR|09HAT8, evidence: FASTV, 71.1%ID,

99.2%length, match=1245)"

/codon_start=1

/protein_id="BAC31913.1"

/db_xref="GI:26336456"

/translation="MVLGNPDVGSPTSDLOHPGSGSCIIISCPGEALAGEPIKY

GELIVGVNGCLASGDKGRSRRLASRPBANGVAPDVWGHISTLVKALSNROCH

SISPTSRSHSVVEYTHDSKDMFOIGSTENMIDPVYDTSGGATGEGASOSTI

SRYACRIICDRRPRTARIYAAGFDASNIPLGERSAKKRTDPGLMDGLTTNGVIMH

PAGGSEDSAPGKWRRTISVCNAYTTIRDSRSNRQKRGKLVNENVTQDSLIDCGAT

LMKRTFGLIRATFLKOLEKORSEANARPOCFVSGSTLAFSPARKTAPDQKQW

YKCGHNVHGWGCRREQPORECEPLRLVGPVPLMGDAAGCLDGPSPHFA

PCGWVSEKTRAYMAQTPRLPHGTNAFHAAFCFGAWLTGELGCVRLIFQSPLD"

ORIGIN

Alignment Scores:

pred. No.: 1,879-153 Length: 1879
Score: 1632.00 Matches: 297
Percent Similarity: 81.58% Conservative: 44
Best Local Similarity: 71.05% Mismatches: 73
Query Match: 71.27% Indels: 4
DB: 3 Gaps: 3

US-10-041-030-4 (1-420) x AK044418 (1-1879)

QY 4 ProG1yG1ng1uG1uH1sCyBa1AProAm1yG1uProVa1yETrYg1uLeuVal 23
DB 194 CCGGGGGAAGAA-----GCACGGCAGCGGAGAGAGCCCATCAAGTATGTTGAATCATC 247
QY 24 ValLeuG1yTrAnG1yAlaLeuProAmG1yAPRArg1yAArg1yLysSerArgpne 43
DB 248 GTTCTGGGCTACAAATGGGTCTCTGGCAATGGAGAGAGCGCGCCGCAAGCCCGCTG 307
QY 44 AlaLeuTy1yAArgProLyBa1aAmG1yVal1yEProSerThrValHsVal1IleSer 63

DB 308 GCACTAGCGCGCGGCACATGCAACGAGAGTGAACCCAGATCTACGACACATCTCC 357
QY 64 ThrProG1nAlaSer1yBa1aIleSerCyLyG1yG1nHsSer1yETrYrThLeu 83
DB 368 AACCACTGGTCTCCAAAGCCCGTGAATACCGAGGCGCACACAGCATCTCATTCACACTG 427
QY 84 SerArgAmG1nThrValValValG1uTyTrH1sAPR1yAPR1yAPR1yAPR1yAPR1y 103
DB 428 TCCCGAGTCACCTACCTACCTATGAGTGCACACACAGCAGCAGCAGCAGCAGCAGCAG 487
QY 104 ValG1yArgSerThrg1uSerPro1IleAPR1yAPR1yThrAPR1yIleSerG1ySer 123
DB 488 ATGGCGGCTCTACTAATAACATGATGACTTTGTGTGAACACACATCTCTGTGTGA 547
QY 124 G1nAPR1yThrAPR1yAlaG1n1IleThrg1nSerThr1IleSerArgpne1aCyArg1le 143
DB 548 GGGGCTACA---GAGGCGCTTCTGCGCAAGTACATCTCCGCTATGCTGCAATTT 604
QY 144 ValCyAPRArgAmG1uProTyTrThAlaArg1lePheAlaAlaG1yPheAPR1ySer 163
DB 605 CTCTGTGACCGCGCGCACCTTATCTGCGCGCATCTGCTGCTGCTGCTGCTGCTGCTG 664
QY 164 LyAPR1yPheLeuG1yG1uLyBa1aAlaLyTrP1yLyAPR1yAPR1yAPR1yAPR1y 183
DB 665 AGCAACATCTTCTTGAAGAGCGGCGCAGCAATGAGAGCTCCGATGCTGTGATGAC 724
QY 184 G1yLeuThrThrAPR1yVal1yLeuVal1yMeCh1sProArg1yG1yPheThrg1uG1uSer 203
DB 725 GGGTTACAAACCAATGGGTTCTGTGTGATGACCCAGCAGGTGGCTTCCGAGACTGT 784
QY 204 G1nProG1yVal1yTrArg1u1IleSerValCyG1yAPR1yVal1yThrLeuArg1uThr 223
DB 785 GCCCGAGGTGTGAGGAGATTTCTGTGCGGGAATGTGACACGTGCGGAGTAGC 844
QY 224 ArgSerAlaG1nG1nArg1yLy1yLeuVal1yG1uSerG1nThrAPR1yAlaG1nAPR1y 243
DB 845 CGATCAGCTCAGAGCGGAGAGGCTGTGTGAAGAAAGTCCAACTCTGCAAGACGCT 904
QY 244 SerLeu1leAPR1yCyG1yAlaThrLeuLeuThrArgThrAlaAPR1yLeuPheH1s 263
DB 905 TGGCTATGACTGTGTGTGGGCGACACTGTGTGGCGCACTCCGCTGGTTCGTGAG 964
QY 264 ThrProThrG1nLyHs1IleG1uAlaLeuArg1nG1u1IleAPR1yAlaArgProG1n 283
DB 965 GCAACCACTAACAACCTAAGGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1024
QY 284 CyAPR1yAlaG1uLeuAmThr1yAPR1yAPR1yAPR1yAPR1yAPR1yAPR1yAPR1y 303
DB 1025 TGGCGGAGGCTTGAACACTTGTGCTTCCGAGCGCAGCCGCTGGCGCAGCAGCAG 1084
QY 304 G1uLy---G1nProThrAlaTyTrLeuSerCyG1yHsValHsG1yTrH1sAPR1y 322
DB 1085 GACAAACAGCAACCCGCTGCTACGTCGCTGTGTGATGATGATGATGATGATGATGATG 1144
QY 323 G1yHsArgSerAPR1yG1uAlaAmG1uArg1uCyAPR1yAPR1yAPR1yAPR1yAPR1y 342
DB 1145 GCGTGGCGGAGGAGCAAGCGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1204
QY 343 ProTyTrValProLeuThrLeuG1yCyG1uAlaG1yPheTyValAPR1yAPR1yPro 362
DB 1205 CCTATGTGCTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1264
QY 363 ThrHsAlaPheThrProCyG1yHsVal1yCySerG1yLySer1aLyETrYrTrSer 382
DB 1265 AGCAGAGCTTTTGACACCTGTGGCAGCAGTCTGTGTGAAGAAGACGCGCGCTAGTGGCT 1324
QY 383 G1n1IleProLeuProH1sG1yThrH1sAlaPheH1sAla1yCyAPR1yPheCyAlaTrp 402
DB 1325 CAGACACCTCTGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCT 1384
QY 403 G1nLeuVal1yG1nG1nAmSer1yLeu1yLeu1yPheG1nG1yPro1IleAPR 420

Db 1385 TGGCTACCGGTGAGCTTGCTGTCTCCGGCTCATTTTCCAGGGGCCACTGGAC 1438

RESULT 8
LOCUS AY408790 1242 bp DNA linear GSS 12-DEC-2003
DEFINITION Pan troglodytes HCM341 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY408790
VERSION AY408790.1 GI:39764758
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.
REFERENCE 1 (bases 1 to 1242)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
AUTHORS 2 (bases 1 to 1242)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
source 1..1242
location/Qualifiers
gene /organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/locus_tag="HCM341"

ORIGIN

Alignment Scores:

| Pred. No.: | 3,56e-153 | Length: | 1242 |
|------------------------|-----------|---------------|------|
| Score: | 1626.50 | Matches: | 298 |
| Percent Similarity: | 80.91% | Conservative: | 41 |
| Best Local Similarity: | 71.12% | Mismatches: | 73 |
| Query Match: | 71.03% | Indels: | 7 |
| DB: | 9 | Gaps: | 4 |

US-10-041-030-4 (1-420) x AY408790 (1-1242)

QY 3 SerProGlyGlnGluGlnHisCySalAProAnLysGlnProValIleTyrGlyGluLeu 22
Db 1 TCTCCCGGT-----GAAGATGGCAGCCAGGCGAGAGCCCATCAAGTATGTTGAATCTC 54
QY 23 ValIleuGlyTyrAsnGlyAlaLeuProAnGlyAspArgGlyAArgArgLysSerArg 42
Db 55 ATGCTCTGGGCTACATAGGAGTGTCTGGCAAGTGGGACAAAGGCGCGCGGAAGCGCC 114
QY 43 PheAlaLeuTyrIleArgProLysAlaAnGlyValIleProSerThrValHisValIle 62
Db 115 CTGGCACTAGCGCGCGCGCGCAAGCGGGGTAAACCCAGAGCTATGCACCAATC 174
QY 63 SerThrProGlnAlaSerLysAlaIleSerCySlyGlyGlnHisSerLieserTyrThr 82
Db 175 TCCAGCGCC-----CTCGCACTAGCACTAACCGGTGTCAACACAGCATCTCGTATACA 225
QY 83 LeuSerArgAnGlnThrValIleGlyTyrThrHisAspLysAspThrAspMetPhe 102
Db 226 CTGTCTCGGAGCACTGGGTCAATAGTGGAGTATACACATGAATAGCCACACACATGTC 285
QY 103 GlnValGlyArgSerThrIleuSerProIleAspPheValIleThrAspThrIleSerGly 122

Db 286 CAGNNNGCCCGCTCCACAGAGAACATGATTGACTTCGTGGTAAAGACACATGCTCMMTGA 345
QY 123 SerGlnAenThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArg 142
Db 346 GGA---GGGGCTGCGAGAGGCCCTTCTGCGCCAGAGACACATCTCTCGCTATGCTGCGCG 402
QY 143 IleValCySeAspArgAnGlnLupProTyrThrIleAlaArgIlePheAlaIleGlyPheAspSer 162
Db 403 ATCTCTGTGACCGCGCGCCACCTTATCTGCGCCATCTATGCGCGCTTGCTTCAATGCG 462
QY 163 SerLysAsnIlePheLeuGlyGlyLysAlaIleAlaIleTyrLysAsnProAspIleHisMet 182
Db 463 TCTAACCAATCTTCTCTTGAGAGAGGACCGGCAATGCGGACCCAGATGCGCTGATG 522
QY 183 ArgGlyLeuThrThrAsnGlyValIleuValIleHisProAlaArgIleGlyPheThrGluGlu 202
Db 523 GATGAGACTGACCACTAAATGAGAGTCTGTGATGACCCCGGACGCGGCTTCTCCAGAGAC 582
QY 203 SerGlnProGlyValITrArgGluIleSerValCysGlyAspValTyrThrLeuArgGlu 222
Db 583 TCAGCCCGGGGTGTGGCGGAGATCTCGGTGTGGGAAATGTGTACATCTGCGGAGAC 642
QY 223 ThrArgSerAlaGlnGlnArgGlyLysLeuValGlySerGluThrAsnValIleuGlnAsp 242
Db 643 AGCCCTCAGCCCGACAGCGGCGGAGAGCTGTGTGAAACAGAGTCCAGCTGTCAGAGAC 702
QY 243 GlySerLeuIleAspLeuCyGlyValIleThrLeuLeuTTrArgThrAlaAspGlyLeuPhe 262
Db 703 GGCTCTTCATCGACCTGTGTGGGCGCACGTGCTGTGGCCACACCGCGCGGGGTGTGTG 762
QY 263 HisThrProThrGlnLysHisIleGlnAlaLeuArgGlnIleAsnAlaAlaArgPro 282
Db 763 CGGGCTCCACACTAGACAACTGAGAGCCAGCGGCGAGAAATGCAAGCGGCGCC 822
QY 283 GlnCyProValIleGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValIle 302
Db 823 CAGTCCCGCGGTGGCTCAGACATCTGCGCTTCCACCCAGCGCGGTGGCGCACAGG 882
QY 303 GlnGluLys---GlnProTTrAlaTyrLeuSerCyGlyHisValHisGlyTyrHisAsn 321
Db 883 CCGGCAAAACAGACCGCTGGGTCTACGTCGCGGCGGACGTCCACAGGCGGTACACAGC 942
QY 322 TrpGlyHisArgSerAspThrGlnAlaAnGlyArgLysCyProMetCysArgThrVal 341
Db 943 TGGGCTGCGCGGAGCGGCGGCGCCAGAGAGCGCAATGCTCTCTGCGGCTTGTG 1002
QY 342 GlyProTyrValProLeuTTrPLeuGlyCyGlnAlaGlyPheTyrValAspAlaGlyPro 361
Db 1003 GGGCTTATGTGCTCTATGCTTGGCCAGAGGCGCGCTGCTGAGCCCTGAGCCCTGGGCG 1062
QY 362 ProThrHisAlaPheThrProCyGlyHisValCysSerGlyLysSerAlaIleTyrTrp 381
Db 1063 CTTACCAATGCTTGTTCACCTTGGCGGCGCACGTGCTCTGAGAAAGATGCCCGCTACTGCG 1122
QY 382 SerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIleCysProPheCysAla 401
Db 1123 GCCAGACACATGCGCCAGGACCCAGATCTTTCATGCGCGCTGCGCTTTTGGCGG 1182
QY 402 ThrGlnLeuValIleGlyGlnLysAsnCyIleLysLeuIlePheGlnGlyProIleAsp 420
Db 1183 GCCTGCTTACTGCGGAGACATGCGGTGCGGCTCATATTTCCAGGCGCGCGTGAT 1229

RESULT 9
LOCUS CR606152 2572 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSDN004Yg19 of Adult Brain of Homo sapiens
(human).
ACCESSION CR606152
VERSION CR606152.1 GI:50486959
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1. (bases 1 to 2572)
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact: Peng Liang Email: filiang@life.techn.com URL: <http://fulllength.invitrogen.com/InvitrogenCorporation1600>
 Faraday Avenue
 2. (bases 1 to 2572)
 Genoscope.
 REFERENCE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 AUTHORS BP 131 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
 JOURNAL - Web : www.genoscope.cns.fr
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 FEATURES
 Location/Qualifiers
 1. 2572
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DN004Y19"
 /issue_type="Adult brain"
 /plasmid="pCMVSPORT_6"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.09e-152 Length: 2572
 Score: 1622.00 Matches: 297
 Percent Similarity: 80.38% Conservative: 43
 Best Local Similarity: 70.21% Mismatches: 75
 Query Match: 70.83% Indels: 8
 DB: 3 Gaps: 4
 US-10-041-030-4 (1-420) x CR606152 (1-2572)
 QY 3 SerProGlyGlnGluGluHisCybAlaProAsnLysGluProValLysTrpGlyGluLeu 22
 DB 199 TCTCCCGGT-----GAAGATGGCGCACCGCGAGAGAGCCCATCAATGTGTGAATCT 252
 QY 23 ValValLeuGlyTrpAnGlyAlaLeuProAnGlyAspArgGlyArgLysSerArg 42
 DB 253 ATGCTCTGGGCTACATGATGTTCTTGGCAAGTGGGACAAGGGCGCGCGCAAGCGCG 312
 QY 43 PheAlaLeuTrpArgProLysAlaAnGlyValLysProSerThrValHisAlaVal 62
 DB 313 CTGGCACTGAGCGCGCGTGCACGCCAGCGGGGTGAGCAACGATCATGACCAATC 372
 QY 63 SerThrPro-----GlnAlaSerLysAlaLysSerCybLysGlyGlnHisSer 78
 DB 373 TCCACCGCGCTGACCTTTTCTCCACACGAGCACTGATTAACCGTGTACGACACG 432
 QY 79 IleSerTrpThrLeuSerArgAnGlnThrValValAlaGluTrpHisAspLysAsp 98
 DB 433 ATCTGTATACACTGTCCTCCGAGCACTCGTATAGTGAATACATGATGAGGAC 492
 QY 99 ThrAspMetPheGlnValGlyArgSerThrGlySerProLysPheValValThrAsp 118
 DB 493 ACAGACATGTTCCAGATGGCCCTCCACAGAGAACATGATTACTTCGGTAAACAC 552
 QY 119 ThrIleSerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArg 138
 DB 553 ACCTCCCTCGAGAG---GGGCGTCCGAGAGGCCCTTCCGACAGACACATCTCCGC 609
 QY 139 PheAlaCybArgGlyLeuValCybAspArgAnGluProTrpThrAlaArgIlePheAla 158
 DB 610 TATGCTGGCGGATCTCTGTGACCGCGGACCACTATCTGCGCGCATCTATGCGCT 669
 QY 159 GlyPheAspSerSerLysAsnIlePheLeuGlyGluLysAlaAlaLysTrpLysAsnPro 178

DB 670 GGGCTTGAGGCTCTTGACCAACATCTTCTTGAGAGCGAGCGCCCAATGGCGAGCCCA 729
 QY 179 AspGlyHisMetAspLysLeuThrThrAnGlyValLeuValMetHisProArgGlyGly 198
 DB 730 GATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 769
 QY 199 PheThrGluGluSerGlnProGlyValAlaTrpArgGluIleSerValCybLysAspVal 218
 DB 790 TTCCTCGAGAGACTCACCGCGGCTGCTGCGGAGATCTCGGCTGTGGGAATGTGTC 849
 QY 219 ThrLeuArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsp 238
 DB 850 ACATTCGGAGACAGCGCTCAGCCAGCGAGCGGAGAGTGTGTGAACAGGTCACAC 909
 QY 239 ValLeuGlnAspLysSerLeuIleAspLeuCybGlyAlaThrLeuLeuTrpArgThrAla 258
 DB 910 GTCTCTCAGAGAGCGGCTCTGATCATGACCTGTGTGTGTGTGTGTGTGTGTGTGTGT 969
 QY 259 AspGlyLeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsn 278
 DB 970 GCGGGGCTGCTGCGGCTCCACACAGCACTGAGGAGCCAGCGGAGAGGCAAAAT 1029
 QY 279 AlaAlaArgProGlnArgProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArg 298
 DB 1030 GCAGCGCGGCGCCAGAGTCCCGGCTGAGCCTGAGCCTGAGCCTGAGCCTGAGCCTGAG 1089
 QY 299 LysGluValValGluGluLys---GlnProTrpAlaTrpLysSerCybGlyHisValHis 317
 DB 1090 GCGCGGACAGCGCGCCAGCAACAGACAGCGCTGAGTCTAGCTCCGCTCGCGGACGTCAT 1149
 QY 318 GlyTrpHisAsnTrpGlyHisArgSerAspThrGluAlaAnGluArgGluCybProMet 337
 DB 1150 GGTCTACACAGCGCTGCGGCTGCGCGGAGCGGCGCGCCAGAGAGCGGATCTCTCTC 1209
 QY 338 CysArgThrValGlyProTrpValProLeuTrpLeuGlyCybGluAlaGlyPheTrpVal 357
 DB 1210 TCGCGCTGTGTGGGCTTATGTGCTCTATGTGCTTGTGCGAGAGCGCGCTGCTGCTG 1269
 QY 358 AspAlaGlyProProThrHisAlaPheThrProCybGlyHisValCybSerGlyLysSer 377
 DB 1270 GACCTCGGCGCGCTGAGCCATGCTTGTGACCTGTGCGGACGTCGCTGCTGAGAGACT 1329
 QY 378 AlaLysTrpTrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaVal 397
 DB 1330 GCGCGCTACTGGGCGCCAGACACACACTGCGCCACGAGCACTTCTTCATGCGCGCTG 1389
 QY 398 ProPheCybAlaThrGlnLeuValGlyGluGlnAsnCybLysLeuLeuPheGlnGly 417
 DB 1390 CCTTTTGGCGGCGCTGCTTACCGGAGACATGCTGTGCTGCTGCTTATTTCCAGGGC 1449
 QY 418 ProIleAsp 420
 DB 1450 CCGCTGAT 1458
 RESULT 10
 AY408789 1242 bp DNA linear GSS 12-DBC-2003
 LOCUS Homo sapiens HCM341 gene, VIRUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY408789
 VERSION AY408789.1 GI:39764757
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1. (bases 1 to 1242)
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejaritwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civejlo, D.R., Lu, F., Murphy, B.,
 Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shtinsky, J.J.,
 Adams, M.D. and Cargill, M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous

gene tritos
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 1242)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Snihsy,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1.1242
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>1242
 /locus_tag="HCM3341"
 gene
 Alignment Scores:
 Pred. No.: 1.14e-152 Length: 1242
 Score: 1621.50 Matches: 297
 Percent Similarity: 80.91% Conservative: 42
 Best Local Similarity: 70.88% Mismatches: 73
 Query Match: 70.81% Indels: 7
 DB: 9 Gaps: 4
 US-10-041-030-4 (1-420) x AY408789 (1-1242)
 QY 3 SerProG1yGlnGluGluHisCysAlaProAsnLysGluProValLysTyrGlyGluLeu 22
 Db 1 TCTCCCGT-----GAAGATGGCGACGCCAGGAGAGCCCATCAAGATGTGTAAC 54
 QY 23 ValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgLysSerArg 42
 Db 55 ATCGTCTGGGCTACATGATGTTGTCTGCGACAGTGGGAGACAGGCGCGCGCAAGCCCG 114
 QY 43 PheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHisValIle 62
 Db 115 CTGGCACTAGCGCGCGTGCACGCCAACGGGTGAACCCAGCATGACCAACCAATC 174
 QY 63 SerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSerTyrThr 82
 Db 175 TCCAGCCCG-----CTCGCAGTGAATACCGGTGTCAGACACGATCTGTATACA 225
 QY 83 LeuSerArgAsnGlnThrValValValGluTyrThrHisAspLysAspThrAspMetPhe 102
 Db 226 CTGTCCCGGAGCCATCGGTGATGAGTATACATGATGATGACACACACATGTC 285
 QY 103 GlnValGlyArgSerThrGlnSerProIleAspPheValValThrAspThrIleSerGly 122
 Db 286 CAGATTTGGCGCTCCACAGAGAAATGATGATGTTGTGTGTAACAGACACGTCCTCGA 345
 QY 123 SerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArg 142
 Db 346 GGA---GGGGTGGCCGAGGGCCCTTGTGCCAGAGACCATCTCCCGCATGCTGCC 402
 QY 143 IleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPheAspSer 162
 Db 403 ATCTCTGTGACCGCGCGCCACTTACTGCGCCGATCATGCGCTGCTGCTGATGCC 462
 QY 163 SerLysAsnIlePheLeuGlyGlyLysAlaAlaLysTyrLysAspProAspGlyHisMet 182
 Db 463 TCTAGCAACATCTCTTGGAGAGCGAGCGCCAAATGCGAGACCCCAATGCGCTGATG 522
 QY 183 AspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyLysPheThrGlnGlu 202
 Db 523 GATGACATGACCAACATGAGTCCCTGTGTATGACACCCCGAGCGGCTTCTCCAGAGC 582
 QY 203 SerGlnProGlyValTyrArgGluIleSerValCysGlyAspValTyrThrLeuArgGlu 222

Db 583 TCAGCCCCGGGTGTCTGGCGGGAGATCTCGGTCTGGGAGATGTATACATTTGGCGGAC 642
 QY 223 ThrArgSerAlaGlnGlnAlaArgGlyLysLeuValGlnSerGluThrAspValLeuGlnAsp 242
 Db 643 ACCCCCTCAGCCACAGAGGGGAGAGCTGTGTGAAGAAAGATCTCAACCTGTGCAAGAC 702
 QY 243 GlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrArgThrAlaAspGlyLeuPhe 262
 Db 703 GGCTCTCATGACACTGTGTGGGCCACACATGCTGTGGCGGACACCGCGGGGCTGTG 762
 QY 263 HisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAspAlaAlaArgPro 282
 Db 763 CGGGCTCCACACTAAGACATGAGAGGCCCGCGGAGAGAAATGACAGCGGGGCC 822
 QY 283 GlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValVal 302
 Db 823 CAGTCCCGCGTGGGCTCTGACACTGTGGCTTCCAGCCAGCCCGGTGGCGCACAGCG 882
 QY 303 GlnGluLys---GlnProThrAlaTyrLeuSerCysGlyHisValHisGlyTyrHisAsn 321
 Db 883 CCCGACAAACAGACACCTCGGTCTACGTCGCTCGGGCACAGTCCATGCTACCAAGCG 942
 QY 322 TyrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThrVal 341
 Db 943 TGGGGCTGGCGGGAGCGGGGCCGCCAGAGCGCGAATGTCTCTGCGCGCTGTG 1002
 QY 342 GlyProTyrValProLeuTyrLeuGlyCysGlyAlaGlyPheTyrValAspAlaGlyPro 361
 Db 1003 GGGGCTTATGTGCTCTATGTGCTTGTGCGACAGGCGGGCTCTGCTGGACCTTGAGCGG 1062
 QY 362 ProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyrTrp 381
 Db 1063 CCTAGCAATGCTTTGACACCTTGGCGGACAGTCTGTGAGAAACATCGCCGCTACTCG 1122
 QY 382 SerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCysAla 401
 Db 1123 GCCCAGACACCACTGCCACCGGACCCATGCTTTCATGCGCGCTGCCCCCTTTGGCGG 1182
 QY 402 ThrGlnLeuValGlyGlnGluAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
 Db 1183 GCCTGACTTACCGGGAGAGATGCTGCGTCCGCTCATATTTCCAGAGGCCGCTGAT 1239
 RESULT 11
 AY408791 1242 bp DNA linear GSS 12-DEC-2003
 LOCUS AY408791
 DEFINITION Mus musculus HCM3341 gene, VIRUAL TRANSCRIPT, partial sequence.
 ACCESSION AY408791
 VERSION AY408791.1 GI:39764759
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1242)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Snihsy,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering

TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 2474)
AUTHORS Adachi, S., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohtsuo, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT A cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
source Location/Qualifiers
1. .2474
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:DB:9230114E15"
/db_xref="taxon:10090"
/clone="9230114E15"
/sex="male"
/tissue_type="epididymis"
/clone_1ib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
<1. .854
/note="unnamed protein product; pellino 2 (MGI:1891445, GB|NM_033602, evidence: BLASTN, 99%, match=864)
putative"
/codon_start=3
/protein_id="BAC28485.1"
/db_xref="GI:26329493"
/translation="RPRCRILVCDRNEPYARIFAAGFSSKXNIFIGEXAKXKPNPGH MGGITNGVIVAHFPGSGFTBESOPGWTREISVCGDYTTTRBTSACQKGLVESSTNV LQDGSILIDCGATLLWRADGLFHAFTQKHEALRQEIINARQPCGVGNTLAPEISIN RKEVEKQPMWATLSCGHVGHYSGWRBDETEENRECEMCRVGVYVPLMIGCEGIF YVDAGPPTTHAFTPCGHVCSSEKAKWSQIPLPHGTIAFHAACPCFATQIVGFONCIKL IFQGPVD"
polyA_signal 2446. .2451
/note="putative"
polyA_site 2474
/note="putative"
ORIGIN
Alignment Scores: 2.65e-146 Length: 2474
Pred. No.: 1563.00 Matches: 279
Score: 99.65% Conservative: 3
Percent Similarity:

Best Local Similarity: 98.59% Mismatches: 1
Query Match: 68.25% Indels: 0
DB: 3 Gaps: 0
US-10-041-030-4 (1-420) x AK033815 (1-2474)
QY 138 ACPheaaIACyAaRgI]eValCySaPaRgaEngIuProTyRThraIaArgIlePheaa 157
DB 3 AGCTTCGATGACGATGCTGTGTGACAGGAACGACCATACAGCAGCATATTCCGG 62
QY 158 AAGIyPheAaPSeSerIySaAnIlePheUeUgIyUySaIaAlaYsTrIySaAn 177
DB 63 GGAGGATTTCGATTCTCCAAAATAATCTTTCTTGAGAGAAAGCAGCAAAATGAAAAAC 122
QY 178 ProAPRGVYHMeAaPBGlyeUeUhrThraEngIyValIleValMeHIsProARgIy 197
DB 123 CCGATGACACATGATGATGACTCTACCAACGGTCTGTGTATGACCCCGAAGGA 182
QY 198 GIyPheThrgIuGIuSerGIuProGIyValTrPaRgIuIleSerValCySgIyAaPvaI 217
DB 183 GGCCTTCACCGAGGAATCCAGCTGAGTCTGAGAGAGATCTCTGTGGGATGTA 242
QY 218 TyRThraUaRgIuThraRgSerIaagIuInaRgIyUySLeUaIgluSerIuThr 237
DB 243 TACACCTTCGAGACACACAGGTGCGCCAGCAGAGGGAAAGCTGTGAAATGAGACC 302
QY 238 AaenVallleuGIaAaPBGlySerIeUleIaPleUeUySgIyValaThreUeUtrPaRgThr 257
DB 303 AACGCTTCGAAAGACGGCTCCTCAATGACCTGTGTGGGCCACCTCTCTGGAGAAC 362
QY 258 ALaAaPBGlyeUeUheHIsThraProThrgIuYHIsIleGuaIaUeUaRgIuInIle 277
DB 363 GCAAGTGGCTTTTTCAGCGCTCTCACTCAAGAACACATAGAACCTCCGAGAGAGATC 422
QY 278 AaenIaAlaAaRgProGInCyBProVaIglYleUaenThraUaPheProSerIleAa 297
DB 422 AATGACAGCCGACCCAGAGTGGCCCTGTAAACACCTGCTTCCCGACATCAAC 482
QY 298 ArgIySgIuValIgluGIuIySgIuProTPaIaTyRLeUSeCySgIyHIsValHIs 317
DB 483 CGAAGGAAGTGTGAAGAGAGAGCGCTGGGATACCTAGCTGCGGACATGTGCAC 542
QY 318 GIYTyRHIsaenTrIglYHIsaRgSerAaPThrgIuaIaenGIuAaRgIuCyBProMet 337
DB 543 GGCTACCAACAGCTGGGGCATCGAGCGACAGAAAGCAAGAGGGAGTGTCCCATG 602
QY 338 CySaGThraValGIyProTyRValProleUtrPleUeUgIyCySgIuAaIglYHIsVal 357
DB 603 TCGAGGACTGTGGGCGCTTACCTCTCTGTGGGCTGTGAGGACAGATTTATATGC 662
QY 358 AaPaIaGIyProProThraHIsaIaPheThraProCySgIyHIsValCySergIuIySser 377
DB 663 GATGGGGACCCCAACTACGCTTACCCCTCGGGCAGCTGTTCAGAAAGTCT 722
QY 378 ALaYsTyRTrSerGIuIleProleUProHIsGIyThraIaIaPheHIsaIaIaCyS 397
DB 723 GCCAAGTACTGTTCGAGATCCACTGCCCGACGAAAGCAGCGCTTACGCCCTGT 782
QY 398 ProPheCySaIaThGIuIeUaIglYgluInaRySgIleUeUleIlePheGInglY 417
DB 783 CGTTCGTGGCCACAGCTGTGTGTGAACAGAACTGCATCAATTTATTTCCAAAGT 842
QY 418 ProIleAaP 420
DB 843 CCAATGAC 851
RESULT 13
BX388547
LOCUS BX388547 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
DEFINITION BX388547 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
CDNA clone CSODD006F18 5-PRIME, mRNA sequence.
ACCESSION BX388547
VERSION BX388547.1 GI:30460950

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3531)

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE Direct Submision

JOURNAL Submitted (16-Jul-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

FEATURES

source

1.3531

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM:DB:4930404N17"

/db_xref="taxon:10090"

/clone="4930404N17"

/sex="male"

/issue_type="testis"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

1.3531

/note="peellino 1 (MGD|MG1:1914495, GB|NM_023324, evidence: BLASTN, 99%, match=2228)"

ORIGIN

misc_feature

misc_feature

Alignment Scores:

Pred. No.: 5,64e-137 Length: 3531

Score: 1473.00 Matches: 271

Percent Similarity: 85.75% Conservative: 42

Best Local Similarity: 74.25% Mismatches: 50

Query Match: 64.32% Indels: 3

DB: 3 Gaps: 0

US-10-041-030-4 (1-420) x AK029586 (1-3531)

QY 58 ThValHieValIleSerThProGlnAlaSerIleAlaIleSerCysIleGlyGlnHis 77

Db 2350 ACTGCAATATTTGCTGCTTAACCTTTCATCTCATAGCAATAGCAACAGACCGCAT 2409

QY 78 SerIleSerTyThrIleuSerArgAnGlnThrValValGlnTyThrHisAspLys 97

Db 2410 ACCATATCATATATCTTATCTCAGGCCAGACGCTGGTGGTAATAATATCTATGACAGC 2469

QY 98 AspThAspMetPheGlnValGlyArgSerThrGlnSer-ProIleAspPheValAla 117

Db 2470 AACCACTGATATGTTTCAAGTGGTCGCTCAACTGAAAGTCGCTATTGATTTGTATGTAAG 2529

QY 117 rAspThrIleSerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSe 137

Db 2530 TCACACCGTTCGTGGAAGTCAGATGTAATTCGACACGAGTACAGTACAAAGACATTTATC 2589

QY 137 rArgPheAlaCysArgIleValCysAspArgAnGlnProTyThrAlaArgIlePheAl 157

Db 2590 AAGATTGGCTGTAGATCATATGTAGAGCGACGTCCTTACAGCTCGAGATTATGC 2649

QY 157 aAlaGlyPheAspSerSerIleValIlePheLeuIleGlyIleValAlaIleTyThrLys 177

Db 2650 TGCAGGGGTGTATCATCAAAAACATCTTTCTTGGGAGAAAGCGTCGCAAGTGAAGAC 2709

QY 177 nProAspGlyHisMetAspGlyLeuThr-Thr-AsnGlyValIleuValMetHisProArg 197

Db 2710 ATGTATGGCGCAGATGATGCTTACACAGCTCAATAGAGTTCTTGTGATGCATCCAGTA 2769

QY 197 lGlyPheThrGlnGlnSerGlnProGlyValTyrArgGlnIleSerValCysGlyAspV 217

Db 2770 ATGGGTTCCAGAAACATCCAAACGTGGAAATGTGAGAAATATCATATGTGGAAACG 2829

QY 217 aITyThrIleuArgGlnThrArgSerAlaGlnGlnArgGlyValLeuValGlnSerGlnT 237

Db 2830 TCTTACGTGTGCTAAAGCCAGATCAGCTCAGCAGAGAAAGTGTGAAATTTGAA 2889

QY 237 hTrAsnValIleuGlnAspGlySerLeuIleAspLeuCysGlyValAlaThrIleuLeuTyArgT 257

Db 2890 CCAATCAGGTACAAATATGGCTCTTAATGATGACCTTTGTGTGCAACCGCTGTGGCGTA 2949

QY 257 hAlaAspGlyLeuPheHisThrProThrGlnLysHisIleGlnAlaIleuArgGlnGlnT 277

Db 2950 CTGCAGAAAGGCTTTCCCATACTCTGCTAGCTGAACACTTAAACCTTAAACAGCGAGA 3009

QY 277 lAsnAlaAlaArgProGlnCysProValGlyLeuAsnThrIlePheProSerIleA 297

Db 3010 TTAAGCGGCTGCGCCAGAGTCCCTGTAGGGTTCAACACATAGCTTCCCAAGATGA 3069

QY 297 sArgLysGlnValValGlnGlnIleProTyAlaTyThrIleuSerCysGlyHisValH 317

Db 3070 AAGAGAGAGATGTTTGAATTAAGTAAAGCAACATGGGTATATTAACGCGGCATGTCC 3129

QY 317 lGlyTyThrHisAsnTrpGlyHisArgSerAspThrGlnAlaAsnGlnArgGlnCysProm 337

Db 3130 ATGCTTATCATTAACGTGGGAAACAAAGAAACGAGACGGCAAAATCGTGAATGCCCTTA 3189

QY 337 eCysArgThrValAlaGlyProTyValProLeuTyPheGlnCysGlnAlaGlyPheTyV 357

Db 3190 TGTGTAGGTCTGTGTGCTTATGTCCTCTGTGGCTTGAATGTGAAGCTGAATTTTATG 3249

QY 357 aAlaAspAlaGlyProProThrHisAlaPheThrProCysGlyHisValCysSerGlnLys 377

Db 3250 TGGACGGCGGCGCTCCACCCATGCTTTAGCCCTGTGGGCAAGTGTTCAGAAAGAC 3309

QY 377 eAlaLysTyThrTySerGlnIlePheProLeuProHisGlyThrHisAlaPheHisAlaIac 397

Db 3310 A-ACGGCTTACGTGTCACAGATCCCATCTCCATGATGATGACGACACTTTTCAGAGCT 3368

QY 397 yAspPheCysAlaThrGlnIleuValGlyGlnGlnAsnCysIleLysLeuIlePheGln 417

Db 3369 GCCCTTCTGTCAATCATGTGGCTGTGTAACAAAGGCTATATCAAGATTATTTCCAG 3428

QY 417 lYProIleAsp 420

Db 3429 GACCTTTAGAC 3439

RESULT 15

CN329888 957 bp mRNA linear EST 14-APR-2004

LOCUS CN329888/c

DEFINITION AGECOURT 22003665 NICHD XCC Te2 xenopus laevis cDNA clone

IMAGE:7207690 3', mRNA sequence.

ACCESSION CN329888

VERSION CN329888.1 GI:46387524

KEYWORDS EST.

SOURCE Xenopus laevis (African clawed frog)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 9, 2004, 13:12:47; Search time 7535 Seconds
(without alignments)
2635.922 Million cell updates/sec

Title: us-10-041-030-4
Perfect score: 2220
Sequence: 1 MFSPQGEHCAPNKEPVKYG.....ATQVGEONCKILIFQGPID 420

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4526729 seqs, 2364849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n model -DEV=x1p
-O=/cgn2.1/USPTO.spool.p/US10041030/runat.09122004.095517.7663/app.query.fasta_1.583
-DB=GenEmbl -QFMT=fasterp -SUPFIX=rgc -MINMATCH=0.1 -LOOPT=0 -LOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10041030 @CCN 1.1 7406 @runat.09122004.095517.7663 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -BSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_hg:.*
3: gb_in:.*
4: gb_m:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_pl:.*
8: gb_pr:.*
9: gb_ro:.*
10: gb_sc:.*
11: gb_sy:.*
12: gb_un:.*
13: gb_vl:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|----|--------------------|
| 1 | 2290 | 100.0 | 1565 | 6 | AX817203 Sequence |
| 2 | 2290 | 100.0 | 1811 | 6 | BC009476 Homo sapi |
| 3 | 2290 | 100.0 | 5579 | 6 | BD190328 Ekl1 phos |
| 4 | 2290 | 100.0 | 5597 | 9 | AF302502 Homo sapi |

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|--------|-------------|-----------|----|-------------|
| 5 | 2290 | 100.0 | 5921 | 6 | AX686405 |
| 6 | 2283 | 99.7 | 1263 | 6 | AR483568 |
| 7 | 2283 | 99.7 | 1263 | 6 | AX298079 |
| 8 | 2175.5 | 95.0 | 1260 | 6 | AR483567 |
| 9 | 2175.5 | 95.0 | 1260 | 6 | AX298077 |
| 10 | 2138.5 | 93.4 | 1646 | 5 | BC072891 |
| 11 | 2129.5 | 93.0 | 1936 | 5 | BC063200 |
| 12 | 2111.5 | 92.2 | 1717 | 6 | BD190327 |
| 13 | 2111.5 | 92.2 | 1717 | 6 | BD190327 |
| 14 | 2006 | 87.6 | 2474 | 5 | BC075973 |
| 15 | 1917 | 83.7 | 1257 | 6 | AR483566 |
| 16 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 17 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 18 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 19 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 20 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 21 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 22 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 23 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 24 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 25 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 26 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 27 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 28 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 29 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 30 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 31 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 32 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 33 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 34 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 35 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 36 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 37 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 38 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 39 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 40 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 41 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 42 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 43 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 44 | 1917 | 83.7 | 1257 | 6 | AX298075 |
| 45 | 1917 | 83.7 | 1257 | 6 | AX298075 |

ALIGNMENTS

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|----|-------------|
| 1 | 2290 | 100.0 | 1565 | 6 | AX817203 |
| 2 | 2290 | 100.0 | 1811 | 6 | BC009476 |
| 3 | 2290 | 100.0 | 5579 | 6 | BD190328 |
| 4 | 2290 | 100.0 | 5597 | 9 | AF302502 |

REFERENCE
AUTHORS
Eisenhardt, G., Kuner, R., Lanahan, A., Worley, P., Spielvogel, D. and
Schek, S.
TITLE
JOURNAL
The P30 gene family and the utilization thereof
Axaton Bioscience AG (DE)
LOCATION/Qualifiers
1. 1565
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
51. 1313
/note="unnamed protein product; codirender Bereich der
M33 cDNA des Menschen"
/codon_start=1
/protein_id="CAB84330.1"

/db xref="GI:39722595"
 /crlinalation="MPSRQGEHCAPNKEPYKGYLVLGYNGALPNDRGRKSRFA
 LYKRPKANGVSKSTYHVISIPQASRKAISCKGHSITYLSNRQIVNEVETHTKDTMP
 QVGRSTESPIDFVVDITLISGQNTDEAQITOSTISRFAIRIVCDNEPETAIFAGF
 DSSKNIPLGEKAKAKWNPQGHNDGLTTNGALVMPHPRGPTBESQGVWEISVCDVY
 TLRETRSAQORGLKVESETNVLQDGLIDLCATLLMRADLFTPTQKLEALROE
 INAAPOCVGLANTLAFPSINRKEVEESQPMAYLSCGHVYHNHMGHSDPEANERE
 CPMCRVGVPLMTIGCEAGFPVDAGPPTHAFTPGHNCSESAKYMSQIPLPHGTHA
 FHAACPFCATQVBEONCKILFQGPID"

ORIGIN

Alignment Scores:

Pred. No.: 3,87e-177 Length: 1565
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.00% Conservatve: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-10-041-030-4 (1-420) x AX817203 (1-1565)

QY 1 MetPheSerProGlyGlnGluGluHscYValAProAenlyGluProValIlyTyGly 20
 Db 51 AGGTTTCCCTGGCCAGAGAGAACACTGCGCCCAATAGAGAGCAAGTGAATACGCG 110
 QY 21 GtuleuValIleuGlyTyraenGlyAlaIleuProAenGlyAspAargIlyArgIlys 40
 Db 111 GAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 170
 QY 41 SerArgPheAlaLeuTyTyraenGlyProIyValAaenGlyValIlyProSerThrValHis 60
 Db 171 AGTAGATTGGCTCTCAACAAGGGGCCCAAGGCAANTGGTCAACCCAGCACCGTCCAT 230
 QY 61 ValIleSerThrProGlnAlaSerIlyValIleSerCysIyGlyGlnHisSerIleSer 80
 Db 231 GTGATATCACCCGCCAGCATCAAGGCTATCAGCTGCAAGAGTCAACACAGTATATCC 290
 QY 81 TyTrPheSerArgAaenGlnThrValIleValIleGluTyTrHisIleAspIyAspThrAsp 100
 Db 291 TACACTTGTGACAGGATACAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 350
 QY 101 MetPheGlnValIyArgSerThrGluSerProIleAspPheValIlyThrAspThrIle 120
 Db 351 AGTTTCAAGTGGGAGATCAACAGAAAGCCTATGACTTGTGTCTCACAGACAGATT 410
 QY 121 SerGlySerGlnAaenThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 411 TCTGCAGCCAGAACACGAGCAGAGCCAGATCACACAGACACCATATCCAGTTCGCC 470
 QY 141 CysArgIleValIyCysAspArgAaenGluProTyTrThrAlaArgIlePheAlaIleGlyPhe 160
 Db 471 TCCAGGATGTGTGGACAGAAATGAACCTTACACAGCAGGATATTCGCCCGCGGATT 530
 QY 161 AspSerSerIyAsenIlePheIleuGlyGluIyValAlaIleTyTrIyAsanProAspGly 180
 Db 531 GACTCTTCCAAAACATATTCTTGGAGAAAAGGAGCAAGTGGAAAAACCCGACGCG 590
 QY 181 HisMetAspGlyLeuThrThrAenGlyValIleuValMetHisProArgIlyGlyPheThr 200
 Db 591 CACATGGATGGGCTCACTATATGCGCTCCGATGATCATCCAGAGGGGCTTCAC 650
 QY 201 GtIuIySerGlnProGlyValIyTrArgGluIleSerValIyCysGlyAspValIyThrIleu 220
 Db 651 GAGGAGTCCACCCCGGGGTCTGGGCGAGATCTCTGTGTGGAGATGTGACCTTG 710
 QY 221 ArgGluThrArgSerAlaGlnIlyValIyLeuValIyGluSerGluThrAsnValIleu 240
 Db 711 CGAGAAACAGGTCGGCCACGACAGAGAAAAGCTGTGGAAAAGTGAACCAACGCTCG 770
 QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaIleThrLeuLeuTrArgThrAlaAspGly 260
 Db 771 CAGGACGGCTCCCTCATTTGACCTGTGTGGGCGACACTCTCTGTGAGAACACAGATGG 830

QY 261 LeuPheHisThrProThrGlnIlyHisIleGluAlaIleuArgGlnGluIleAsnAlaIa 280
 Db 831 CTTTTCATACCTCCACACTAGAGACATAGAGCCCTCCGACAGATTTAACCGCGCC 890
 QY 281 ArgProGlnCysProValIyIyLeuAaenThrIleuAlaIlePheProSerIleAsnArgIyGlu 300
 Db 891 CGGCTCATGTGCTCTGTGGGGGTCAACACCCGCTTCCCGACATCAACAGAAAG 950
 QY 301 ValValIyGluIyGluProIyTrAlaIyIleuSerCysGlyValIleValIleGlyTyHis 320
 Db 951 GTGGTGGAGAGAAACACCCCTGGGCAATATCTCACTTGTGGCAGCTGACGCGTACAC 1010
 QY 321 AenTrpGlyHisArgSerAspThrGluAlaAaenGluArgGluCysProMetCysArgThr 340
 Db 1011 AACTGGGGCCATCGAGTGAACAGAGGCCAAGAGAGAGTGTCCATGTGCGAGACT 1070
 QY 341 ValGlyProTyValIyProIyTrPheIyGlyCysGluAlaIyPheTyValIleAspAlaGly 360
 Db 1071 GTGGGCCCTATGTCTCTCTGTGGCTGTGGAGAGATTTATATGACACCGACGA 1130
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValIyCysSerGluIySerAlaIyTyTr 380
 Db 1131 CGGCCACTCATGTCTTCACTTCCCTGTGACACGCTGTCTCGAAGAGCTGCAAAATAC 1190
 QY 381 TrpSerGlnIleProIyProIyTrHisIyHisAlaPheHisAlaIyCysProPheCys 400
 Db 1191 TGGTCTCAGATCCCTTGTCTTCACTGAGACTCATGATTTACGCTGTGCTTCCCTTCTGT 1250
 QY 401 AlaThrGlnLeuValIyGluGlnAaenCysIleIyLeuIlePheGlnIyProIleAsp 420
 Db 1251 GCTACACAGCTGTGTTGGGAGAGCAAACTGATCAATTAATTTCCAGAGTCAATTGAC 1310

RESULT 2
 BC009476 1811 bp mRNA linear PRI 29-JUN-2004
 LOCUS Homo sapiens pellino homolog 2 (Psoeph11a), mRNA (cDNA clone
 DEFINITION MG:15066 IMAGE:3942712), complete cds.
 ACCESSION BC009476
 VERSION BC009476.2 GI:33872093
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1811)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Murnane, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Uebelin, T.B., Tosliyluk, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McMan, P.J.,
 McKernan, K.J., Malek, J.A., Gunatane, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, S.W.,
 Villalón, D.K., Mizny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Scherchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywicki, M.I., Skalka, U., Smilun, D.E.,
 Scherchen, A., Schein, J.E., Jones, S.J. and Marra, M.A.,
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 JOURNAL 12477932
 PUBMED 2 (bases 1 to 1811)
 REFERENCE Strausberg, R.
 DIRECT SUBMISSION Submitted (25-JUN-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK

USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:14550456.
 Contact: MGC help desk
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadan@systemsbio.org
 Anup Madan, Jessica Pahey, Erin Helton, Mark Kerteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL plate: 25 Row: 9 Column: 3
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10864062.
 Location/Qualifiers

FEATURES

source

1.1811
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:15066 IMAGE:3942712"
 /issue_type="Lung, small cell carcinoma"
 /clone_id="NIH-MGC-7"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"

gene

CDS

1.1811
 /gene="PEL12"
 /db_xref="LocustID:57161"
 287..1549
 /gene="PEL12"
 /codon_start=1
 /product="pellino 2"
 /protein_id="AAH09476.1"
 /db_xref="GI:14550457"
 /db_xref="LocustID:57161"
 /translation="MPSQGEHCAKPNKPKVKGSLVLYNGALPNDGRGRKSRFA
 LYKRPANGVKSTVIVISIPQSKAISGKQHSISYLSRQTVVVEVTHDQDMF
 QVGRSTPSPIDPVDTISGQNTDEAQTOSTISFARIQDREPEPTATIPAGF
 DSKNIFLEKAKAKMKNPQGHMDGLTTNGVLVHMRPGSTESQPEVMEISGQVY
 TURETSAQQRGLVSEETNVLDGSLIDLCATLLMRADGLFHTPKHLEAKRE
 INAAEQCEVGNATLLAFPSINKEVVEKQPMAYLSCGVHGVHNMGHMSDEANRE
 CPMCRVGVPLMLGCEAGFYVDAGPPTHAFTPCGVHSEKSAKYWSQIPPHGTHA
 FHACPFCAATQLVGEQNCIKLIFQSPID"

ORIGIN

Alignment Scores:

Pred. No.: 4.62e-177 Length: 1811
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-041-030-4 (1-420) x BC009476 (1-1811)

QY 1 MetPhSerProGlyGlnGluGluHsiGyeAlaProAsnLysGluProValLysTrgLy 20
 DB 287 ATGTTTCCCTCGGCGAGGAGAACTGCGCCCAATTAAGAGCCAGTGAATACGGG 346
 QY 21 GluLeuValAlaLeuGlyTrAenGlyAlaLeuProAsnGlyAspArgGlyATrGATG 40
 DB 347 GAGCTGGTGGTCTCGGGTACATGCTTACCCCATGAGATGAGAGACGAGGAAA 406
 QY 41 SerArgPheAlaLeuTrLysATrGProLysAlaAsnGlyValLysProSerThrValHis 60
 DB 407 AGTAGATTGGCCCTCTCAACAGCGCCCAAGCAAAATGGTGTCAAAACCCACGCTCCAT 466
 QY 61 ValILSerThrProGlnAlaSerLysAlaILSerCyeAlaGlyGlnHisSerLys 80
 |||||

DB 467 GTGATATCCAGCCCGGAGCATCCAGGCTATCAGCTCGCAAGGTCAACAGATATATCC 526
 QY 81 TyrThrLeuSerArgAsnGlnThrValValValGluTrpHisAspLysAspThrAsp 100
 DB 527 TACACTTTGTCAAGAAATACAGACTGTGGTGGAGTACACACATGATTAAGATACGAT 586
 QY 101 MetPheGlnValGlyATrSerThrTrpLysProLysAspPheValThrAspThrLe 120
 DB 587 AGTTTCAGTGGTGGAGATCAACAGAAAGCCCTATCGACTTCGTGTCAACACAGATT 646
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrLysSerArgPheAla 140
 DB 647 TTGGCAGCGCAACACGAGACGAAACCCAGATCAACAGAGCAATATCCAGTTCGCC 706
 QY 141 CysArgGlyValCysAspArgAsnGluProTrpThrAlaArgLysPheAlaGlyPhe 160
 DB 707 TCCAGATGTGTGCGACAGAAATGAACCTTACACAGACGGAATATTCGCCCGGATTT 766
 QY 161 AspSerSerLysAsnLysPheLeuGlyGluValAlaLysTrpLysAspProAspGly 180
 DB 767 GACTCTTCCMAAAACATATTTCTTGGAGAAAGGACGAAAGTGAATAAAACCCCAAGCC 826
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 DB 827 CACATGATGGGCTCACTACTAATGCGCTCGTGATGCATCCACAGAGGGGCTTCAAC 886
 QY 201 GluGluSerGlnProGlyValITrArgGluLysSerValCysGlyAspValTrpThrLeu 220
 DB 887 GAGGAGTCCAGCCCGGGGTGTGGCGAGATCTGCTGTGGAGATGTGTACACTTGG 946
 QY 221 ArgGluThrArgSerArgGlnGlnIleArgLysValLeuValGluSerGluThrAsnValLeu 240
 DB 947 CGAGAAACCGAGTGGCGCCAGCAAGAGAAAGCTGGTGAAGAGTGAAGCCAAAGCTTGG 1006
 QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTrpArgThrAlaAspGly 260
 DB 1007 CAGGAGCGTCTCTCATTAACCTGTGTGGGCGCACTCTCTTGAAGAACACAGATGGG 1066
 QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaAla 280
 DB 1067 CTTTTTCATACCTCAACTCAGAAAGCATATGAAGCCCTCCGACAGATTAACCCCGCC 1126
 QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerLysAsnArgLysGlu 300
 DB 1127 CGGCTCACTGCTCTGTGGGCTCAACCCCTGGCTTCCAGATCAACAGAAAGAG 1186
 QY 301 ValValGluGluLysGlnProTrpAlaTrpLysSerCysGlyHisValHisGlyTrpHis 320
 DB 1187 GTGTGGAGAGAAAGAGCCCTGGGCAATTCATGTTGTGGCAGCTGACCGGTAACCC 1246
 QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProPheCysArgThr 340
 DB 1247 AACTGGGGCATCGAGGTGACACGAGGCGCAAGAGGAGTGTCCCATGTGCGAGACT 1306
 QY 341 ValGlyProTrpValProLeuTrpLeuGlyCysGluAlaGlyPheTrpValAspLysGly 360
 DB 1307 GTGGGCGCTTAATGTGCTCTGTGGCTGTGAGGACGAGATTATGTATACCGAGGA 1366
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTrp 380
 DB 1367 CCGCCACATCATGCTTTCATCTCCCTGTGACACGCTGTCTCGGAAAGCTCCAAAATTC 1426
 QY 381 TrpSerGlnIleProLeuProHisGlyTrpHisAlaPheHisAlaAlaCysProPheCys 400
 DB 1427 TGTCTCAGATCCCGTGTGCTCATGAACTCATGATTTCAAGCTGCTTGGCCCTTCTCT 1486
 QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysValLysValLeuIlePheGlnGlyProLysAsp 420
 DB 1487 GCTACACAGCTGTGGGAGCAAAATGCATCAAAATTAATTTTCCAAAGGTCCAAATTCAC 1546
 RESULT 3
 BD190328
 LOCUS BD190328 5579 bp DNA linear PART 17-JUL-2003

DEFINITION Elkl1 phosphorylation related gene.
ACCESSION BD190328
VERSION BD190328.1 GI:33000067
KEYWORDS MO 03008589-A/68.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Matsuzaki, O., Matsuda, A., Nagano, Y. and Suzuki, N.
Elkl1 phosphorylation related gene
Patent: WO 03008589-A 68 30-JAN-2003;
JOURNAL ASAHI KASEI CORP, OSAMU MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI

COMMENT OS Homo sapiens (human)
PN MO 03008589-A/68
PD 30-JAN-2003
PF 15-JUL-2002 MO 2002JP007174
PR 18-JUL-2001 JP 01P 218204,31-AUG-2001 JP 01P 263450 PR
21-JAN-2002 JP 02P 012176
PI OSAMU MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI PC
C12N15/54, C12N9/12, G01N33/50, G01N33/15, C07K16/40, A61K39/385, PC
A61K31/711

FEATURES
source Location/Qualifiers
1..5579
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1..8e-176 Length: 5579
Score: 2290.00 Matches: 420
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-041-030-4 (1-420) x BD190328 (1-5579)

QY 1 MetPheSerProGlyGlnGlnGluHISGYSAlaProAsnLysGluProValLysTyrGly 20
DB 177 AAGTTTCCCTTGGCCAGAGGAAACCTGGCCCCCAATAGAGCCAGTGAATAACGGG 236

QY 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgLys 40
DB 237 GAGCTGGTGGTCTCGGGTACAAATGCTTACCAATGGAATAGAGAGAGAGAGAA 296

QY 41 SerArgPheAlaLeuTyrTyrAspGlyPheAlaAsnGlyValLysProSerThrValHis 60
DB 297 AGTAGATTGGCCCTTCAACAGCGGCCCAAGGCAAAATGCTGCAAAACCCGACCCGTCAT 356

QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
DB 357 GGTATTCACGCCCCAGGCACTCCAGGCTATCAGCTGCAAGGTCACACACAGTATATCC 416

QY 81 TyrThrLeuSerArgAsnGlnThrValValAlaGluTyrThrHisAspLysAspThrAsp 100
DB 417 TACACTTGTCAAGAAATCAGACTGCGTGGTGAATACACATCATATAAGATACGAT 476

QY 101 MetPheGlnValAlaArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
DB 477 AGGTTTCAGGTGGGCGATCAACAGAAACCCCTATGACTCTGCTGTCAACAGACGATT 536

QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 537 TTTCGACGCAACACCGAGCAAGCCCGATCACAACAGACACCATATCCAGGTTCCGC 596

QY 141 CysArgIleValAlaCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160

DB 597 TGCAGATGCTGTGGACAGAAATGAACTTACACAGACGGATATTGCGCGGGAATTT 656

QY 161 AppSerSerLysAsnIlePheLeuGlyGlyLysAlaAlaLysTyrLysAsnProAspGly 180
DB 657 GACTCTCCAAATAATATTTCTTGAGAAAAGGACGAAAGTGAATAAAACCCCAACGCG 716

QY 181 HisMetLysGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
DB 717 CACATGATGGCTCTACTATATGAGCTCTGGTGAATGATCCAGAGGGGCTTCAC 776

QY 201 GluGluSerGlnProGlyValTyrPargLysIleSerValCysGlyAspValTyrThrLeu 220
DB 777 GAGGATGCCACCGCGGGGTCTGGCGGAGATCTGTCTGTGAGAGATGTATACCTTG 836

QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
DB 837 CGAGAAACAGGTCGCGCCAGCAGCAAGAGAAAGCTGAGAAAGTGAAGCCAACTCTCG 896

QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrPargThrAlaAspGly 260
DB 897 CAGACGGCTCCCTATTTGACCTGTGTGGGCTCCTCTCTGGAAGAACAGATGGG 956

QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaAla 280
DB 957 CTTTTCATATCTCCAACTCAGAAACATPAGAAGCCCTCCGCGAGAGATTAAACGCCGCC 1016

QY 281 ArgProGlnCysProValAlaGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
DB 1017 CGGCTCAGGTCTCTGTGGGCTCAACACCTGGCTTCCCAAGCATCAACGAGAAAG 1076

QY 301 ValValGluGluLysGlnProThrPalaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
DB 1077 GTGTGAGAGAACACAGCCCTGGGCAATATCTCACTTTGGCCACGTCAACGGGTACAC 1136

QY 321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGluArgLysCysProMetCysArgThr 340
DB 1137 AACTGGGGCCATCGAGTGAACCGGACCAACGAGAGAGTGTCCCATGTGACGACT 1196

QY 341 ValGlyProTyrValProLeuThrLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
DB 1197 GTGGGCCCTTATGTCTCTCTGTGGCTGTGGTGGAGAGATTTATGTAGAACCCAGAGA 1256

QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
DB 1257 CGGCCAATCTCATCTTCACTCTCTGTGACACGTTGTCTGGAGAAATCTGCAGAAATTC 1316

QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
DB 1317 TGGTCTCAGATCCGTTGCTCATATGAACTCATGATTTCAAGCTTGCCTTCTGT 1376

QY 401 AlaThrGlnLeuValAlaGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
DB 1377 GCTACACAGCTGTGTGGGAGCAAAATGATCAATTAATTTTCCAAAGTCCATTGAC 1436

RESULT 4
AF302502 5597 bp mRNA linear PRI 05-JUL-2001
LOCUS Homo sapiens pellino 2 (PELI2) mRNA, complete cds.
ACCESSION AF302502
VERSION AF302502.1 GI:10242352
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Reisch, K., Jockusch, H. and Schmitt-John, T.
Assignment of homologous genes, Pel11/PELI1 and Pel12/PELI2, for
the Pelle adaptor protein Pellino to mouse chromosomes 11 and 14
and human chromosomes 2p13.3 and 14q21, respectively, by physical
and radiation hybrid mapping
JOURNAL Cytogenet. Cell Genet. 92 (1-2), 172-174 (2001)
MEDLINE 21203570

PUBMED 11306823
2 (bases 1 to 5597)
REFERENCE Resch, K., Jockusch, H. and Schmitt-John, T.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (05-SEP-2000) Developmental Biology and Molecular
Pathology, University of Bielefeld, Universitaetstrasse 25,
Bielefeld, NRW 33615, Germany

FEATURES
Location/Qualifiers

source 1..5597

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

1..5597

/gene="PEL12"

177..1439

/gene="PEL12"

/note="similar to Drosophila melanogaster Pellino"

/codon_start=1

/product="pellino 2"

/protein_id="AA015390.1"

/db_xref="GI:10242353"

/translation="MPSQOEHCAPNKEPVKXGLVILGYNALPNDGRKSRFA
LYKRPKANGVKESTVITSPQSKAISCQHSISTYLSRNTVAVETHDQTMF
OYGRSTESPIDPVYDTISGQNTDEAQTOSTISFACRYDREPEYTAIRFAGP
DSKRI FLBKAKAKMPDGMHDLTTNGVVMHPRGCTBESQPVMBEISGQVY
TKETRSADQKGLVESETNVLDGSLIDCGATILMRADGLFHTPTKHI EALRQ
INARPOCPVGLNTLA FPSINRKEVVEKPPAYLSGCHVHGYNHGRHSDEANRE
CPMCRITGVPEVPLMLGCEAGFYVDAGPPTTAFPPCGHVCSEKAKYWSQI PLPHGTHA
FHACPFCAATQVGEQNCIKLIRQPID"

ORIGIN

Alignment Scores:

Pred. No.: 1..8e-176 Length: 5597
Score: 2290.00 Matches: 420
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-041-030-4 (1-420) x AF302502 (1-5597)

QY 1 MetPheSerProGluGlnGluGlnH1sCYeAlaProAsnLysGluProValLysTyrGly 20
DB 177 ATGTTTCCCTCGGCGAGGAGAACACTGCGCCCAATTAAGAGCGAGGAATACGGG 236
QY 21 GluLeuValValLeuGlyTyrAengLYALaLeuProAsnLYAspArgLYATrGly 40
DB 237 GAGCTGTGTGCTCGGATCAATGCTTACCCCAATGAGATAGAGACGAGAGAA 296
QY 41 SerArgPheAlaLeuTyrLysArgProLYAspAlaAsnGlyValLysProSerThrValHis 60
DB 297 AGTAGATTGCTCTTCAAGAGCGCCCAAGGCAAAATGATGTCAACCCAGACCGTCAT 356
QY 61 ValLysSerThrProGlnAlaSerLysValLysSerCysLysGlyGlnHisSerLysSer 80
DB 357 GTGATATCCAGCGCCCAAGGATCCAAAGGCTATACGCTGCAAGAGTCAACACATATATCC 416
QY 81 TyrThrLysSerArgAsnGlnThrValValValGlyTyrThrHisAspLysAspThrAsp 100
DB 417 TACACTTTGTCAAGAAATCAGACTGTGTGTGTGAGATACACATGATTAAGATACGAT 476
QY 101 MetPheGlnValGlyArgSerThrGluSerProLysPheValValThrAspThrTle 120
DB 477 ATGTTTCAGGTGGCGAGATCAACAGAAAGCCCTATCACTTCCTGTGCACAGACAGATT 536
QY 121 SerGlySerGlnSerThrAspGluValGlnLysThrGlnSerThrLysSerArgPheAla 140
DB 537 TCTGGCAGCGCAGAACCGGAGGAGCCAGATACACAGAGCCATATCCAGATTCGCC 596
QY 141 CysArgLysValCysAspArgAsnGluProTyrThrAlaArgLysPheAlaLysGlyPhe 160
DB 597 TGCAGATGCTGTGCGCAGAGATGAACCTTACACAGCAAGATATTCGCCGCCGATTT 656

QY 161 AspSerSerLysAsnLysPheLeuGlyGlyLysValAlaLysTyrLysAsnProAspGly 180
DB 657 GACTCTTCCMAAAACATATTTCTTGAGAGAAAGGCGAAAGGAGAAAAACCCGACGCC 716
QY 181 HisMetArgGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
DB 717 CACATGATGGGCTCACTATATGAGCTCTGTGTATCATTCACAGAGGGGCTTCACC 776

QY 201 GluLysSerGlnProGlyValTyrArgLysLysSerValCysGlyAspValTyrThrLeu 220
DB 777 GAGAGATCCACGCGCGGCTGTGGCGAGATCTGTGTGTGAGATGTGTACACTTG 836

QY 221 ArgGluThrArgSerAlaGlnGlyLysLysValGluSerGluThrAsnValLeu 240
DB 837 CGAGAAACGAGTCCGCGCCAGCAACAGAAAGCTGTGTGAAAGTGTAGCCAACGCTCG 896

QY 241 GlnAspGlySerLeuLysPheLysCysGlyValTyrLeuThrLysThrAlaAspGly 260
DB 897 CAGAGCGCTCCCTCATTAACCTGTGTGGGCCACTCTCTTCGAGAACAGCAGATGGG 956

QY 261 LeuPheHisThrProThrGlnLysHisLysGlyAlaLeuValArgGlnGluLysAsnAla 280
DB 957 CTTTTTCATACCTCAACTCAGAAAGCAGATAGAACCTTCGCGAGAGATTAAACGCCGC 1016

QY 281 ArgProGlnCysProValGlyLysAsnThrLysAlaPheProSerLysAsnArgLysGlu 300
DB 1017 CGGCTCACTGTCTCTGTGGGCTCAACACCTCGGCTTCGCCAGATCAACAGAAAGG 1076

QY 301 ValValGlnGlyLysGlnProTyrAlaTyrLysSerCysGlyHisValHisGlyTyrHis 320
DB 1077 GTGTGTGAG 1136

QY 321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGlyArgGlyCysProMetCysArgThr 340
DB 1137 AACTGGGCGCATCGAGTGACACGAGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1196

QY 341 ValGlyProTyrValProLeuThrLeuGlyCysGlyValGlyPheTyrValAspAlaGly 360
DB 1197 GTGGGCGCTATGTGCTCTGTGCTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1256

QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlyLysSerAlaLysTyr 380
DB 1257 CGGCCACATCATGCTTCACTCTCTGTGAGACAGTGTCTCGAGAGAGCTCGAAATAC 1316

QY 381 TrpSerGlnLysProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
DB 1317 TGGTCAATCCCGTGTGCTCATGGAATCAATGATTTACAGCTGCTTCCCTTCTGT 1376

QY 401 AlaThrGlnLeuValGlyGlnGlnAsnCysLysLysLeuLysPheGlnGlnLysProLys 420
DB 1377 GCTACACAGCTGTGGGAGAGCAAACTGCATCAAAATTAATTTTCCMAAGTCCAAATTC 1436

RESULT 5

AX686405 5921 bp DNA linear PAT 29-MAR-2003

LOCUS Sequence 3 from Patent WO02059611.

AX686405

DEFINITION AX686405.1 GI:29372143

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

Tularik Inc

CDS
/db xref="taxon:9606"
501..1763
/note="unnamed protein product; human pellino 2"
/codon_start=1
/protein_id="CAD86576.1"
/db_xref="GI:2937214"

/translation="MPSQGEHCAPNKEPVKYGELVVLGYNGALPNGDRGRKRFALYKRPKANGKSTVHISTPOASKASICKQGHISYLTSLKQVAVVEYTHDKDTDMFQVGRSTESPIDFVITDTISGQNTDEAOITOSTISRFACRIYCDNEPYTARIPAGPDSKNITPLGEKAAKKNPGRHNDGTTNNVLMHPRGRTESQGWREISVCCDVTRETRSAQQRGLVSESTNVLDQSLIDLCATLLMRTADGLFHTPOKHEALRQEIINAARPCQCVGLNTLAPSLINKVEVEEKQPAWYLSCGHVGHYHMGHSDTEANBRECPMCRTPGVPLMLTGCSEAGFYVDAGPPHAFPTPGHVCSEKSAKWSQIPPHQSTHAFHACPCFATQVGEONCIKLFQGPID"

ORIGIN

Alignment Scores:

Pred. No.: 1,93e-176 Length: 5921
Score: 2290.00 Matches: 420
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-041-030-4 (1-420) x AK686405 (1-5921)

QY 1 MetPheSerProGluGluGluGluHisCysAlaProAsnLysGluProValLysTyrGly 20
DB 501 ATGTTTCCCTGGCGAGGAGAACACTGCGCCCAATTAAGGAGCAAGTAATACGGG 560
QY 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgLys 40
DB 561 GAGCGTGTGTCTCGGTACAAATGTTGCTTTACCAATGAGATGAGAGACGAGGAAA 620
QY 41 SerArgPheAlaLeuTyrTyrAspArgProLysAlaAsnGlyValLysProSerThrValHis 60
DB 621 AGTAGATTGCGCTTCAACAGCGGCCCAAGGCAAAATGGTGCMAACCCAGCACCGCCAT 680
QY 61 ValLysSerThrProGlnLysSerLysAlaLysSerCysValSerGlyGlnHisSerLysSer 80
DB 681 GTGATATCCAGCCCGCCAGGATCCCAAGGCTTATCAGCTGCAGAAAGGTCAACACGATATATCC 740
QY 81 TyrThrLeuSerArgAsnGlnThrValValAlaGluTyrThrHisAspLysAspThrAsp 100
DB 741 TACACTTGTCAAGAGATCAGACTGTGTTGTTGAGATACACATGATTAAGATACGAT 800
QY 101 MetPheGlnValGlyArgSerThrGluSerProLysAspPheValValThrAspThrIle 120
DB 801 ATGTTTCAGGTGGCGAGATCAACAGAAACCCCTATCGACTTGTGTCTCAGACACGATT 860
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrLysSerArgPheAla 140
DB 861 TCTGGACGCCAAGACCGAGACGAGCCAGATCACACAGCACCCATATCCAGTTTCGCC 920
QY 141 CysArgLysValCysAspArgAsnGluProTyrThrAlaArgLysPheAlaAlaGlyPhe 160
DB 921 TCGAGGATCTGTGCGACAGAGATGAACCTTACACAGCACGAGTATTCGCCCGCGATT 980
QY 161 AspSerSerLysAlaAsnLysPheLeuGlyGlyValAlaAlaLysTyrLysAsnProAspGly 180
DB 981 GACTCTTCCAAACAACTATTTCTTGAGAAAGCGCAAGGAGGAAAAACCCCGACGCGC 1040
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
DB 1041 CACATGATGGGCTCACTAATGAGCGTCTGTTATGATCATCCACGAGGGGCTTCACCC 1100
QY 201 GluGluSerGlnProGlyValTyrArgGluLysSerValCysGlyAspValTyrThrLeu 220
DB 1101 GAGAGATCCCAAGCCCGGGGTCTGGCGAGATCTCTGTGAGAGATGTATACACTTG 1160
QY 221 ArgGluThrArgSerLysGlnGlnArgGlyValLeuValGluSerGluThrAsnValLeu 240
DB 1161 CGAGAAACCAAGGTGGCGCCAGCAACGAGAAAGCTGTGTGAAAAGTGAGACCAACGTCTCG 1220

QY 241 GlnAspGlySerLeuIleAspLeuCysGlyValAlaThrLeuLeuTyrArgThrAlaAspGly 260
DB 1221 CAGGAGGCTCCCTATTGACCTGTGTGGGCCACTCTCTCTGGAGAACAGCAGATGGG 1280
QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaAla 280
DB 1281 CTTTTCATCTCCAACTCAGAAAGCAGATAGAGCCCTCCGGCAGAGATTAACCCGCC 1340
QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
DB 1341 CGGCTTCAGTGTCTGTGGGCTCAACACCTGTGCTTCCAGCATCAACAGGAAAG 1400
QY 301 ValValGluGluLysGluProThrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
DB 1401 GTGGTGGAGAGAACAGACCCCTGGGCAATCTCACTTGTGGCCAGTCGACGGATACC 1460
QY 321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
DB 1461 AACTGGGGCATCGGAGTGACACGGAGGCCAAGAGAGAGTGTCCCATGTGCAGACT 1520
QY 341 ValGlyProTyrValProLeuThrLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
DB 1521 GTGGGCCCCCTATGTGCTCTGTGCTGTGGTGGAGGAGATTTATGTAGACGACAGA 1580
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
DB 1581 CCGCAACATCATGCTTTTCACTCCCTGTGGACACGTTGCTCGGAGAACTCGCAAAATAC 1640
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
DB 1641 TGTCTCAGATCCCTTGTCTCATATGAACTATGATTTTCAACGCTTGTGCTTCTGT 1700
QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProLysAsp 420
DB 1701 GCTACACAGCTGTGTGGAGCAAACTGCATCAATTAATTTTCAAGGTCCAAATGAC 1760

RESULT 6
AR483568 1263 bp DNA linear PAT 14-MAY-2004
LOCUS AR483568
DEFINITION Sequence 7 from patent US 6703487.
ACCESSION AR483568
VERSION AR483568.1 GI:4724652
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1263)
AUTHORS Bird,T.A. and Cosman,D.J.
TITLE Human pellino polypeptides
JOURNAL Patent: US 6703487-A 7 09-MAR-2004;
FEATURES Location/Qualifiers
1..1263
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 1.11e-176 Length: 1263
Score: 2283.00 Matches: 419
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 1
Query Match: 99.69% Indels: 0
DB: Gaps: 0

US-10-041-030-4 (1-420) x AR483568 (1-1263)

QY 1 MetPheSerProGluGluGluGluHisCysAlaProAsnLysGluProValLysTyrGly 20
DB 1 ATGTTTCCCTGGCGAGGAGAACACTGCGCCCAATTAAGGAGCAAGTAATACGGG 60
QY 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgLys 40

Db 61 GACCTGCTGCTCGGATCAATGCTCTTACCCATGAGATGAGAGACGAGAAA 120
 Qy 41 SerArgPheAlaLeuTyrLybArpProLybAlaAsnGlyValLybProSerThrValHis 60
 Db 121 AGTAGATTGGCTCTTCAACAGGGCCCAAGGCAAAATGGTGTCAAAACCCAGCACGTCAT 180
 Qy 61 ValIleSerThrProGlnAlaSerLybAlaIleSerCybLybGlyGlnHisSerIleSer 80
 Db 181 GTGATATCCAGCCGCCAGGCATCAAGGCTATCAGCTGCAGAAAGTCAACACATATATATCC 240
 Qy 81 TyrThrLeuSerArgbAnGlnThrValValValGlyTyrThrHisAspLybAspThrAsp 100
 Db 241 TACACTTTGTCAAGAAATCAGACTGTGTGTGTGAGTACACATGATTAAGATTACGAT 300
 Qy 101 MetPheGlnValGlyArgSerThrGlySerProIleAspPheValValThrAspThrIle 120
 Db 301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGACCTCTGTTCACACAGACGATT 360
 Qy 121 SerGlySerGlnAsnThrAspGlyAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 361 TCTGGCAGCCAGAACACGAGCAAGCCAGATCAACAGAGCACCATATCCAGGTTCCGCC 420
 Qy 141 CysArgIleValCysAspArgbAnGlnProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
 Db 421 TGCAGATCGTGTGCGACAGAAATGAACCTTACACAGACGAAATTCGCCGCCGATTT 480
 Qy 161 AspSerSerLybAsnIlePheLeuGlyGlyLybAlaAlaIleTyrLybAsnProAspGly 180
 Db 481 GACTCTTCCAAAACATATTTCTTGAGTAAAGCCAGCAAGTGGAAAACCCCGACGCG 540
 Qy 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgIlyGlyPheThr 200
 Db 541 CACATGATGGCTCACTACTAATAGCGCTCTGTGTGATGATCAACAGGGGCTTCACC 600
 Qy 201 GluGlnSerGlnProGlyValIleTPAArgGlnIleSerValCysGlyAspValIleThrLeu 220
 Db 601 GAGAGATCCAGCCCGGGGTCTGCGGAGATCTGTGTGTGAGATGTGTACACCTTG 660
 Qy 221 ArgGlnThrArgSerAlaGlnIleArgGlyLybLeuValGlnSerGlnThrAsnValLeu 240
 Db 661 CGAGAAACACAGCTCGGCCACAGCAAGAAAGCTGTGGAAGTGAACCAACGTCCTG 720
 Qy 241 GlnAspGlySerLeuIleAspLeuCybGlyAlaThrLeuLeuTyrArgThrAlaAspGly 260
 Db 721 CAGAGAGGCTCCCTCATTTGACTGTGTGTGGGCCACTCTCTGAGAAACAGCAGATGGG 780
 Qy 261 LeuPheHisThrProThrGlnLybHisIleGlnAlaLeuArgGlnGlnIleAsnAlaAla 280
 Db 781 CTTTTTCATATCTCAACTCAGAAACATAGAAAGCCCTCCGCGAGGATTAAACGCCGCC 840
 Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLybGln 300
 Db 841 CGGCTCATAGTCTCTGTGGGGCTCAACACCTGCGCTCCCGACATCAACAGAAAGAG 900
 Qy 301 ValValGlnGlnLybGlnProTyrAlaTyrLeuSerCybGlyHisValHisGlyTyrHis 320
 Db 901 GTGGTGAAGAGAGAGCAGCCCTGGGCTATCTCAGTTGTGGCCACGTCAGCGGTATCAC 960
 Qy 321 AsnTyrGlyHisArgSerArgbAnGlnAlaAsnGlnArgGlyCysProMetCybArgThr 340
 Db 961 AACGTGGGCGCATGGAGTGAACAGAGGCAACAGAGGAGTGTCCCATGTCCAGAGACT 1020
 Qy 341 ValGlyProTyrValProLeuTyrPLeuGlyCysGlnAlaGlyPheTyrValAlaAspAlaGly 360
 Db 1021 GTGGGCCCTATAGTCTCTCTGTGCTGTGAGGACAGATTTATATAGACGACGAGA 1080
 Qy 361 ProProThrHisAlaPheThrProCybGlyHisValCysSerGlyLybSerAlaLybTyr 380
 Db 1081 CCGCAACTCATGCTTTCACCTCCCTGTGACACGTGTGTCGAGAGAGTCTGCAAAATAC 1140
 Qy 381 ThrSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
 Db 1141 TGGTCTCAGATCCCGTGTGCTCATGAGAACTATGCAATTTCAACGCTGCTCCCTCTTGT 1200

Qy 401 AlaThrGlnLeuValGlyGlnGlnAsnCybIleLybLeuIlePheGlnGlyProIleAsp 420
 Db 1201 GCTACACAGCTGTGTGGGAGCAAAATGTCATCAATATTTTCCAAAGTCCAAATGAC 1260
 RESULT 7
 AX298079
 LOCUS AX298079 1263 bp DNA linear PAT 26-NOV-2001
 DEFINITION Sequence 7 from Patent WO0183739.
 ACCESSION AX298079
 VERSION AX298079.1 GI:17128160
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1
 AUTHORS Bird, T. A. and Cosman, D. J.
 TITLE Human peptidic polypeptides
 JOURNAL Patent: WO 0183739-A 7 08-NOV-2001;
 IMMUNEX CORPORATION (US)
 FEATURES
 source 1..1263
 location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ALIGNMENT Scores:
 Pred. No.: 1, 11e-176 length: 1263
 Score: 2283.00 Matches: 419
 Percent Similarity: 99.76% Conserved: 0
 Best Local Similarity: 99.76% Mismatches: 1
 Query Match: 99.69% Indels: 0
 Gaps: 0
 DB: 6
 US-10-041-030-4 (1-420) x AX298079 (1-1263)
 Qy 1 MetPheSerProGlyGlnGlnGlnHisCybAlaProAsnLybGlnProValIlyTyrGly 20
 Db 1 ATGTTTCCCTCGGCCAGAGAGAACCTGCGCCCAATAGAGGCACTGAAATACGCGG 60
 Qy 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGlyb 40
 Db 61 GAGCTGTGTGCTCGGATCAATGATGTGCTTAAACCAATGAGATAGAGAGAGAGAAA 120
 Qy 41 SerArgPheAlaLeuTyrLybArpProLybAlaAsnGlyValLybProSerThrValHis 60
 Db 121 AGTAGATTGGCTCTTCAACAGGGCCCAAGGCAAAATGGTGTCAAAACCCAGCACGTCAT 180
 Qy 61 ValIleSerThrProGlnAlaSerLybAlaIleSerCybLybGlyGlnHisSerIleSer 80
 Db 181 GTGATATCCAGCCGCCAGGCATCAAGGCTATCAGCTGCAGAAAGTCAACACATATATCC 240
 Qy 81 TyrThrLeuSerArgbAnGlnThrValValValGlyTyrThrHisAspLybAspThrAsp 100
 Db 241 TACACTTTGTCAAGAAATCAGACTGTGTGTGTGAGTACACATGATTAAGATTACGAT 300
 Qy 101 MetPheGlnValGlyArgSerThrGlySerProIleAspPheValValThrAspThrIle 120
 Db 301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGACCTCTGTTCACACAGACGATT 360
 Qy 121 SerGlySerGlnAsnThrAspGlyAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 361 TCTGGCAGCCAGAACACGAGCAAGCCAGATCAACAGAGCACCATATCCAGGTTCCGCC 420
 Qy 141 CysArgIleValCysAspArgbAnGlnProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
 Db 421 TGCAGATCGTGTGCGACAGAAATGAACCTTACACAGACGAAATTCGCCGCCGATTT 480
 Qy 161 AspSerSerLybAsnIlePheLeuGlyGlyLybAlaAlaIleTyrLybAsnProAspGly 180
 Db 481 GACTCTTCCAAAACATATTTCTTGAGTAAAGCCAGCAAGTGGAAAACCCCGACGCG 540

QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 DB 541 CACATGAGATGGGCTCACTACTAATGGCGCTCTGTATGATCATCAGAGGGGCTTCACG 600
 QY 201 GluGluSerGlnProGlyValValTParGlyLileSerValCyseGlyAspValTyThrLeu 220
 DB 601 GAGGATCCACGCCCGGGGCTGTGGCGAGATCTGTGTGTGAGATGTGTACCTTG 660
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGlnThrAsnValLeu 240
 DB 661 CGAGAAACCAAGTCGGCCACAGACAGAAAGCTGTGAAAGTAAAGCAACGCTCTG 720
 QY 241 GlnAspGlySerLeuLileAspLeuCyseGlyAlaThrLeuThrParGlyThrAlaAspGly 260
 DB 721 CAGGAGGGCTCCCTCATTTGACCTGTGTGGGGCCACTCTCTCGAGAAACAGACATGGG 780
 QY 261 LeuPheHisThrProThrGlnLysHisAlleGluAlaLeuArgGlnGluLileAsnAlaAla 280
 DB 781 CTTTTCATATCTCACTCACTCAGAAAGCAATPAAAGCCCTCCGGCAGAGATTAAACGCCGC 840
 QY 281 ArgProGlnCyseProValGlyLeuAsnThrLeuAlaPheProSerLileAsnArgLysGlu 300
 DB 841 CGGCTCTAGTGTCTGTGGGGCTCAACACCTGGCTCTCCCGCATCAACAGAAAGAG 900
 QY 301 ValValGluGluLysGlnProTParAlaTyThrLeuSerCyseGlyHisValHisGlyTyThrHis 320
 DB 901 GTGGTGAAGAGAAAGACGCTTGGGCAATCTCAAGTTGTGGCCAGCTGCACGGGTACAC 960
 QY 321 AsnTParGlyHisArgSerAspThrGluAlaAsnGluArgGluCyseProMetCyseArgThr 340
 DB 961 AACTGGGGCCATCGAGTGCACGAGAGCCCAAGAGAGAGAGTGTCCCATGTGCAGAGACT 1020
 QY 341 ValGlyProTParValProLeuThrLeuGlyCyseGluAlaGlyPheTyThrValAspAlaGly 360
 DB 1021 GTGGGCCCCATATGGCTCTGTGGCTGTGGTGAAGAGATTTATGTAGACGACAGA 1080
 QY 361 ProProThrHisAlaPheThrProCyseGlyHisValCyseSerGluLysSerAlaLysTyThr 380
 DB 1081 CGGCCAATCATCTCTTCACTCCCTGTGAGCAGTGTGTGGGAGAAAGCTGTGAAAATTC 1140
 QY 381 ThrSerGlnLileProLeuProHisGlyThrHisAlaPheHisAlaAlaCyseProPheCyse 400
 DB 1141 TGGTCTCAGATCCCGTTGGCTCATGAACTCATGCAATTCACCGCTCTGCTTCTGT 1200
 QY 401 AlaThrGluLeuValGlyGluGlnAsnCyseLileValLeuLilePheGlnGlyProLileAsp 420
 DB 1201 GCTTACACGCTGTGGGAGCAAACTGCATCAAAATTAATTTCCAAAGTCCCAATTGAC 1260
 RESULT 8
 AR483567
 LOCUS AR483567 1260 bp DNA linear PART 14-MAY-2004
 DEFINITION Sequence 5 from patent US 6703487.
 ACCESSION AR483567
 VERSION AR483567.1 GI:47246651
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclassified.
 1. (bases 1 to 1260)
 AUTHORS Bird, T.A. and Cosman, D.J.
 TITLE Human pellino polypeptides
 JOURNAL Patent: US 6703487-A 5 09-MAR-2004;
 FEATURES Location/Qualifiers
 source 1..1260
 /organism="unknown"
 /mol_type="genomic DNA"
 Alignment Scores:
 Pred. No.: 6,44e-168 Length: 1260
 Score: 2175.50 Matches: 399
 Percent Similarity: 97.38% Conservative: 10

Best Local Similarity: 95.00% Mismatches: 10
 Query Match: 95.00% Indels: 1
 DB: 6 Gaps: 1
 US-10-041-030-4 (1-420) x AR483567 (1-1260)
 QY 1 MetPheSerProGlyGlnGluGlnLysHisCyseValaProAsnLysGluProValLysTyThrGly 20
 DB 1 ATGTTTCCCGGGCCAGAGAGAACCCAGCGGCCCCCAACAGAGAACCCGGTGAATATACGGG 60
 QY 21 GluLeuValValLeuGlyTyThrAsnGlyAlaLeuProAsnGlyAspArgGlyValArgArgLys 40
 DB 61 GAGCTGTGTGTCTCTGGGTGTACAAATGTGTCTTTAATGTGTGAACAGGGGACGAGAGAA 120
 QY 41 SerArgPheAlaLeuTyThrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
 DB 121 AGCAAGTTTGGCTCTTATAGAGGAGCTTACGCCAATGTGTTCMAACCCAGACAAATCCAC 180
 QY 61 ValLileSerThrProGlnAlaSerLysAlaLileSerCyseLysGlyGlnHisSerLileSer 80
 DB 181 ATGTCCTCACACCAAGGCGTCCAAAGGCATCAGCTCCAGAGGACATCAGACATATCG 240
 QY 81 TyThrThrLeuSerArgAsnGlnThrValValAlaGluTyThrHisAspLysAspThrAsp 100
 DB 241 TACACGTTGTCCAGGACGACGAGCGGTAGTGTGAGTGCACACACAGATPAAAGACCGAGC 300
 QY 101 MetPheGlnValGlyArgSerThrGluSerProLileAspPheValValThrAspThrLile 120
 DB 301 ATGTTTCAGTGTGGGAGGTCAACAGAAAGCCCATTTGACTTGTGTGTACAGACAGGTT 360
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnLileThrGlnSerThrLileSerArgPheAla 140
 DB 361 TCCGGCGGTCAAGAACGAAGAT--GCCCATGATCACAGAGACCATCTGTAGTTCGCA 417
 QY 141 CyseArgLileValCyseAspArgAsnLysProTyThrAlaArgLilePheAlaAlaLysPhe 160
 DB 418 TCCAGAGATCGTGTGAGACAGAACGAGCCATATACAGACGATATTCGCGGACAGAAATTC 477
 QY 161 AspSerSerLysAsnLilePheLeuGlyGluLysAlaAlaLysTyThrLysAsnProAspGly 180
 DB 478 GATTTCTCCAAATAATCTTTCTTGTGAGAGAAAGCAACAAATGAAAACCTCTATGGA 537
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 DB 538 CACATGATGAGACTCATTACCAATGATGTCTTGAATGACACCCGACAGAGAGCTTCAC 597
 QY 201 GluGluSerGlnProGlyValValTParGlyLileSerValCyseGlyAspValTyThrLeu 220
 DB 598 GAGGAATCCCAAGCTGTGAGTGTGAGAGAGATCTGTCTGTGGGATGTGTACACCTTG 657
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
 DB 658 CAGAGAACCAAGTCCGCCACAGAGAGGAGAAAGCTGTGTGAAGAGTGAACCAACCTCTG 717
 QY 241 GlnAspGlySerLeuLileAspLeuCyseGlyAlaThrLeuLeuTParArgThrAlaAspGly 260
 DB 718 CAAGACGGCTCCCTATTGACCTGTGGGGCCACTCTCTGTGAGAAACCCAGATGGC 777
 QY 261 LeuPheHisThrProThrGlnLysHisAlleGluAlaLeuArgGlnGluLileAsnAlaAla 280
 DB 778 CTTTTCACAGCTCTTACTCAGAAAGCAATPAAAGCCCTCCGACAGAAATCATACAGCC 837
 QY 281 ArgProGlnCyseProValGlyLeuAsnThrLeuAlaPheProSerLileAsnArgLysGlu 300
 DB 838 CAGCCCAAGTCCCGTGGGCTTAAACACCTCGGCTTCCCAAGATCAACCGGAAGAA 897
 QY 301 ValValGluGluLysGlnProTParAlaTyThrLeuSerCyseGlyHisValHisGlyTyThrHis 320
 DB 898 GTGGTGAAGAGAACAGCCCTGGGCAATCTGAGACTCGGSCATGTGCACGAGTCCAC 957
 QY 321 AsnTParGlyHisArgSerAspThrGluAlaAsnGluArgGluCyseProMetCyseArgThr 340
 DB 958 AACTGGGGCCATCGAGAGCAGCGGAAACCAAGAGAGAGTGTCCCATGTGCAGAGACT 1017

QY 341 ValGlyProGlyValProLeuTTPLeuGlyCyseGluAaglyPheTyValaAspAlaGly 360
 Db 1018 GTGGGCCCCCTACGCTCTCTCTGCGGTGGTGGAGCAGGATTTTATGTCGATGCGGGA 1077
 QY 361 ProProthHhAlaPheThrProCyseGlyVhIvalCyseSerGlyuysSerAlaIvalTyT 380
 Db 1078 CCCCCCACTACAGCTTTTACCCCTGCGGGGACGTGTTCAAGAAAGTCGCGCAAGTAC 1137
 QY 381 TTPSerGlnIleProLeuProHhIaglyThhIvalaPheHhIvalaIaCyseProPheCyse 400
 Db 1138 TGGTCGAGATCCACCTACGCCCCACGGAACGAGCGTTTCATGCGCGCTGCTGTCG 1197
 QY 401 AlaThrGlnLeuValGlyGluGlnAserCyseIleValLeuIlePheGlnGlyProIleAsp 420
 Db 1198 GCCACGACGTGGTGGTGAACAGAACTGCATCAATTAATTTCCAAGGTCGACGTGAC 1257

RESULT 9
 AX298077 1260 bp DNA linear PAT 26-NOV-2001
 LOCUS Sequence 5 from Patent WO0183739.
 DEFINITION AX298077
 ACCESSION AX298077 GI:17128159
 VERSION AX298077.1 GI:17128159
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 Bird, T.A. and Cosman, D.J.
 Human pellino polypeptides
 Patent: WO 0183739-A 5 08-NOV-2001;
 IMMUNEX CORPORATION (US)

FEATURES
 source 1..1260
 Location/Qualifiers
 1..1260
 /organism="Mus musculus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10090"

ORIGIN

Alignment Scores:
 Pred. No.: 6 44e-168 length: 1260
 Score: 2175.50 Matches: 399
 Percent Similarity: 97.38% Conservative: 10
 Best Local Similarity: 95.00% Mismatches: 10
 Query Match: 95.00% Indels: 1
 Gaps: 1

US-10-041-030-4 (1-420) x AX298077 (1-1260)

QY 121 SerGlySerGlnAserThraPheGluAaglyIleThhGlnSerThrIleSerArgPheAla 140
 Db 361 TCCGGGCGGTGACAGAACAGAT---GCCAGATCACACAGACCACTCTAGTTCGCA 417
 QY 141 CyseArgIleValCyseAspArgAsnGlyProTyThhAlaArgIlePheAlaIaGlyPhe 160
 Db 418 TGCAGGATGCTGTGTACAGAAACAGACCATATACAGACGCAATATTCGGGACGATTC 477
 QY 161 AspSerSerIleAsnIlePheLeuGlyGlyIleValaIaIaIaIaIaIaIaIaIaIaIaIaIa 180
 Db 478 GATTTCTTCAAAATATCTTTCTTGAAGAAAGACGAAATGAAAAAACCCTGATGGA 537
 QY 181 HhMetAspGlyLeuThrThraAsnGlyValLeuValMetHhIleProArgGlyGlyPheThr 200
 Db 538 CACATGATGAGACTCACTACATCAATGATGCTGATGATGATGATGATGATGATGATGATG 597
 QY 201 GluGlySerGlnProGlyValTPArgGlyIleSerValCyseGlyAspValTyThhLeu 220
 Db 598 GAGGAATCCAGGCTGAGCTGAGAGAGATCTGTCTGTGGGATGTATCACCTTG 657
 QY 221 ArgGlyThraSerSerAlaGlnGlnArgGlyIleValLeuValGlySerGlyThraAsnValLeu 240
 Db 658 CGAGAGACAGAGTCCGCGCCACAGAGAGGAGAGCTGTGGAAGTACAGACGATCTTG 717
 QY 241 GlnAspGlySerIleuIleAspLeuCyseGlyAlaThhLeuLeuTPArgThraIaAspGly 260
 Db 718 CAAGACGGCTCCCTCAATGACCTGTGGGGCCACTCTCTGAGAAACCGCAGATGGC 777
 QY 261 LeuPheHhIleThrProThrGlnIleVhIleGluAlaLeuArgGlnIleValaIaIa 280
 Db 778 CTTTTCAGCTCTCTACTAGAAACATATAGAAAGCCCTCGGAGGAGATCAATGACGCC 837
 QY 281 ArgProGlnCyseProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgIleGly 300
 Db 838 CGACCCCAAGTCCCGTGGGCTTTACACCTGCGCTTCCACGATCAACCGGAAGAA 897
 QY 301 ValValGluGluIleValProTPAlaTyLeuSerCyseGlyHhIaIaIaIaIaIaIaIaIa 320
 Db 898 GTGGTGAAG 957
 QY 321 AsnTPGlyVhIaArgSerAspThrGluAlaAsnGluArgGlyCyseProMetCyseAspGly 340
 Db 958 AGCTGGGGCCATCGAGGAG 1017
 QY 341 ValGlyProTyValProLeuTTPLeuGlyCyseGluAaglyPheTyValaAspAlaGly 360
 Db 1018 GTGGGCCCCCTACGCTCTCTCTGCGGTGGTGGAGCAGGATTTTATGTCGATGCGGGA 1077
 QY 361 ProProthHhAlaPheThrProCyseGlyVhIvalCyseSerGlyuysSerAlaIvalTyT 380
 Db 1078 CCCCCCACTACAGCTTTTACCCCTGCGGGGACGTGTTCAAGAAAGTCGCGCAAGTAC 1137
 QY 381 TTPSerGlnIleProLeuProHhIaglyThhIvalaPheHhIvalaIaCyseProPheCyse 400
 Db 1138 TGGTCGAGATCCACCTACGCCCCACGGAACGAGCGTTTCATGCGCGCTGCTGTCG 1197
 QY 401 AlaThrGlnLeuValGlyGluGlnAserCyseIleValLeuIlePheGlnGlyProIleAsp 420
 Db 1198 GCCACGACGTGGTGGTGAACAGAACTGCATCAATTAATTTCCAAGGTCGACGTGAC 1257

RESULT 10
 BC072891 1646 bp mRNA linear VRT 03-AUG-2004
 LOCUS Xenopus laevis MGC80329 protein, mRNA (cDNA clone MGC:80329
 DEFINITION IMAGE:5074048), complete cds.
 ACCESSION BC072891 GI:49119163
 VERSION BC072891.1 GI:49119163
 KEYWORDS MGC.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

Xenopodinae; Xenopus; Xenopus.

REFERENCE 1 (bases 1 to 1646)
AUTHORS Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W. and Richardson,P.
TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED 12454917
AUTHORS

2 (bases 1 to 1646)
JOURNAL Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D., Altshuler,S.F., Zeeberg,B., Bucow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heich,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stappleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Ueda,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loggiano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Hellon,E., Kettleman,M., Madan,A., Rodigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E., Scheraga,A., Schein,J.F., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
AUTHORS Klein,S. and Gerhardt,D.S.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
REMARK NIH-WGC Project
COMMENT Contact: XGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Igor Dawid
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@cgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butlerfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Ohi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liles Prabh, Parvaneh Saeedi, Jif Santos, Angélique Scherch, Ursula Skalska, Duane Smallus, Jeff Scott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRK Plate: 152 Row: h Column: 6
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.
Location/Qualifiers
1. 1646
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:80329 IMAGE:5074048"
/tissue_type="Ovary, Xenopus"
/clone_id="NICHD_XGC_Ov1"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1. 1646
gene

/gene="MGC80329"
/db_xref="locusID:443954"
336.1595
/gene="MGC80329"
/product="MGC80329 protein"
/protein_id="AAH72891.1"
/db_xref="GI:4919164"
/db_xref="locusID:443954"
/translation="MPSPSQEBHCAPSKPEVKYGEVLVVLGNGCLPNDRGRKSRFA
LFRKRSKGVNRSVHVISTQASAKSRQHSISFLSRQTVVAYTHDKDMMF
QIGRSTSEIDFVVDTGVSQNDSTQGISFACRIVCDNRPPTARFPAAGFV
SKNIFLEKAKAMPDGHGDTGTVLWHPKGGTDSRPMVEISVCGVYT
LRETSADQKGVNENITLQDSGLVDCATILMTARIGLHTPTOKHLEAFOEI
NNAAPQCVGANTLAFPSINKDVVEDQPMAYIKGTVHGHVHNGHSSTDEANREC
PMKRTIGYPVPLMCEAGSFYVDGPPTHAFPCGHVCSKSAKTSQIPLPHGTHAF
HAACFPAICLIQIGBESCVKLIFQGVVD"

ORIGIN

Alignment Scores:

| Pred. No.: | 9,24e-165 | Length: | 1646 |
|------------------------|-----------|---------------|------|
| Score: | 2138.50 | Matches: | 383 |
| Percent Similarity: | 96.90% | Conservative: | 24 |
| Best Local Similarity: | 91.19% | Mismatches: | 12 |
| Query Match: | 93.38% | Indels: | 1 |
| DB: | 5 | Gaps: | 1 |

US-10-041-030-4 (1-420) x BC072891 (1-1646)

QY 1 MetPheSerProGlyGlnGluGlnIscYsAlaProaenLyGluProValIserTyArgly 20
|||||
Db ATGTTTCCCGCAGTCAGAGGAGACATTCGCCGCCAGTAAGAGACCCGTAAATACGGA 395

QY 21 GluLeuValValIleuGlyTyArgnGlyAlaLeuProaenGlyAspArglyArgArgly 40
|||||
Db GAGCTGGTGCTGCTGGGAGTCAACCGCTGTCTCCCAATGGGAGATCGAGAGGCGGAAA 455

QY 41 SerArgPheAlaLeuTyArglyArgProLyAlaAsnGlyValIleProSerThyValHis 60
|||||
Db AGCAGATTTCACCTTTCAACGCCCCANATTTACGGGGTGAAGCCAGCCAGCTTTCAAT 515

QY 61 ValIleSerThProGlnAlaSerLyAlaIleSerCyLyGlyGlnHisSerIleSer 80
|||||
Db GTGATTTCACCCGCGACAGCTTCCAAAGCAATCAGACAGAGGCGCAACAGCATCTCA 575

QY 81 TyrThrLeuSerArgAsnGlnThrValValValGluTyThrHisAspLyAspThrAsp 100
|||||
Db TTCACATTATCCAGAAAGTCAGACTGTGTGGTAATACACCCATGACACAGACACAGAT 635

QY 101 MetPheGlnValAlaArgSerThrGluSerProIleAspPheValIValThrAspThrIle 120
|||||
Db ATGTTTCAGATTGGAAAGTCAAGAAAGCCCATTTGACTTTGTGTAATCATACGCTT 695

QY 636 ArgTTCAGATTGGAAAGTCAAGAAAGCCCATTTGACTTTGTGTAATCATACGCTT 695

QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
|||||
Db TCTGGCAACCAAGAAC---GATGAACACAGATCACAAAGTACATTTCAAGATTGCA 752

QY 141 CysAlaGlnIleValCysAspArgAsnGluProTyThrAlaAlaGlnIlePheAlaIleAla 160
|||||
Db TGTAGAAATAGTGTGGAATAGAAACCCGCGTACAGAGTAAATTTGCTGCGACAGATT 812

QY 161 AppSerSerLyAsnIlePheLeuGlyGluLyAsnAlaAlaLyTrpLyAsnProAspGly 180
|||||
Db GACTCTCCAAAATAATATTTCTTGTGTAGAAAGCGCCAAAGTGAAGAATCTTCACGCT 872

QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
|||||
Db CATATGACACGACCTAATCAACATGGGTCTGTGATGATCAACAAAGGAGGAGTTTACT 932

QY 201 GluGluSerGlnProGlyValIleTrpArgGluIleSerValCysGlyAspValTyThrLeu 220
|||||
Db GATGAATCAGCGCTGTGTGTGTGGGAGAAATTCGTGCTGTGAGATGTTTATACATTA 992

| QY | 221 | ArgGluThrArgSerAaIGlnGlnArgGlyLeuValaIGluSerGluThrAspValaLeu | 240 |
|------------|--|---|-----------------------------|
| Db | 993 | AGAGAAACCAAGATCTGCTCCACAAAGAGCAACTTGTGAAATGAACAATATTCCT | 1055 |
| QY | 241 | GlnaSpGlySerLeuLeuIleAspLeuCySGlyAlaThrLeuLeuThrArgThrAlaSpGly | 260 |
| Db | 1053 | CAGATAGGTCTCTTGTTGACTTATGTGTGGCCGACCTCTTTGGCGAAACGGCGATGG | 1112 |
| QY | 261 | LeuPheHsThrProThrGlnLeuHsIleIleGluAlaLeuArgGlnGluIleAsnAla | 280 |
| Db | 1113 | TTGCTCCATACGGCCCACTCAAAAACAATTGAAGCTTTGACACAGAAATAAATCTGCC | 1172 |
| QY | 281 | ArgProGlnCySProValaGluLeuAsnThrLeuAlaPheProSerTlaAsnArgGlyLeu | 300 |
| Db | 1173 | AGGCTCAAGTGCCTGGTGGAATTAACACTTTAGCTTTCCAGTACCAACCGCAAAAT | 1233 |
| QY | 301 | ValValaGluGluValaGlnProThrAlaTyrLeuSerCySGlyHsValaHsIleGlyThrHs | 320 |
| Db | 1233 | GTTGTAGAAAGCAAAACACGCTTTGGGCTTACCTCAAGTGTGTGCATGTGCATGTGATCAT | 1292 |
| QY | 321 | AsnThrGlyHsIleArgSerAspThrGluAlaAsnGluArgGluCySPromCyAsnArgThr | 340 |
| Db | 1293 | AACTGGGAGACATCGCAGTGCACAGAAACCAAGCAACCGAGTGTCCATGTGTAGACA | 1355 |
| QY | 341 | ValGlyProTyrValaProLeuThrLeuGlyCySGluAlaGlyPheTyrValaAspAlaGly | 360 |
| Db | 1353 | ATTGGTCCCTATGTGCTCTCTATGTGGTGTGTGTGAAGAGATTTTATGTGCATGTGCGA | 1412 |
| QY | 361 | ProProThrHsAlaPheThrProCySGlyHsValaCySerGluLySerAlaTyr | 380 |
| Db | 1413 | CCTCTACTACAGCTTTCAAGCTTTGTGACACGCTGCTCTGAAAAATTCGCCAAATAT | 1472 |
| QY | 381 | TyrSerGlnIleProLeuProHsGlyThrHsAlaPheHsAlaAlaCySPheProHsCyS | 400 |
| Db | 1473 | TGTCGCCAGATTCGCTCCCAATGTGTACACAGCGGTTTACAGCGCTTGCCTTCTGT | 1532 |
| QY | 401 | AlaThrGlnLeuValaGlyGluGlnAsnCySleIleLeuIlePheGlnGlyProIleAsp | 420 |
| Db | 1533 | GCAATATCAATTGACTGTGGGAGAGAGCTGTGTCAAACTTATTTTCAAGGCGCTTGTGAC | 1592 |
| RESULT 11 | | | |
| LOCUS | BC063200 | 1996 bp | mRNA linear VRT 30-JUN-2004 |
| DEFINITION | Xenopus tropicalis hypothetical protein MGC75650, mRNA (CDNA clone | | |
| ACCESSION | BC063200 | | |
| VERSION | BC063200.1 | | |
| KEYWORDS | MG. | | |
| SOURCE | Xenopus tropicalis (Silurana tropicalis) | | |
| ORGANISM | Xenopus tropicalis | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| AUTHORS | Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; | | |
| | Xenopodine; Xenopus; Silurana. | | |
| | 1 (bases 1 to 1936) | | |
| | Strausberg, R.L., Collins, E.A., Grouse, L.H., Derge, J.G., | | |
| | Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., | | |
| | Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K., | | |
| | Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., | | |
| | Datchenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L., | | |
| | Stetpleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., | | |
| | Schnee, T.E., Brownstein, M.J., Ubed, T.B., Toshitsuki, S., | | |
| | Carrinci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., | | |
| | Abraham, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., | | |
| | McKernan, K.C., Malek, J.A., Gunaratne, P.H., Richards, S., | | |
| | Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huly, S.W., | | |
| | Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibby, R.A., | | |
| | Falley, J., Helton, E., Kettleman, M., Madan, A., Rodriguez, S., | | |
| | Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., | | |
| | Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., | | |
| | Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., | | |
| | Butterfield, Y.S., Krzywiec, M.I., Skalka, U., Smalls, D.E., | | |
| | Schmetcher, A., Schein, J.E., Jones, S.J., and Marra, M.A. | | |
| | Generation and initial analysis of more than 15,000 full-length | | |
| TITLE | | | |

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

human and mouse cDNA sequences.
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16699-16903 (2002)
12477932
2 (bases 1 to 1936)
Klein,S. and Strausberg,R.
Direct Submission
Submitted (02-BEC-2003) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: cgapubs-remail.nih.gov
Tissue Procurement: Robert M. Grainger
CNA Library Preparation: Life Technologies, Inc.
CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Buterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Glin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Tatka Olson, Diana Palmquist, Anna Petrescu, Anna Lisa Prabhu,
Parvaneh Saeedi, JR Santos, Angeliqne Scherch, Ursula Skalska,
Diane Smalls, Jeff Scott, Miranda Tsai, George Yang, Jacque
Schein, Asim Siddiqui, Rob Holt, Marco Matra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/URL at: <http://image.lnl.gov>
Series: IRAX plate: 142 Row: m Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.
Location/Qualifiers

```

/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="cazon:8364"
/clone="MGC:75650 IMAGE:5379633"
/tissue_type="Embryo, Siliurana tropicalis, tailbud (stages
20-27)"
/clone_1fb="NICHD_XGC_Emb7"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
1..1936
/gene="MGC75650"
/db_xref="LocusID:394778"
365..1624
CDS

```

ORIGIN

[illegible]

US-10-041-030-4 (1-420) x BC063200 (1-1936)

QY 1 MetPheSerProGlyGlnGlnGlnIleTCysAlaProAlenlySGlnProValIleSTyrGly 20
 Db 365 ATGTTTCCCCAGACAGAGAGACATGCGCCGCCAGTAGAGACCCCTCAAAATACGAA 424
 QY 21 GluLeuValIleLeuGlyTyrAsnGlyAlaLeuProAlenlyAAspArgGlyArgArglys 40
 Db 425 GAGCTGGTGCTGCTGGGGTACAAACGGCTGTCTCCAAATGGGGATCGAGGAAGCGCGAAA 484
 QY 41 SerArgPheAlaLeuTyrIleArgProIlyAlaAsnGlyValIleAspSerThrValHis 60
 Db 485 AGCAGATTGGCACTCTTAAACGCCCTTAATCCAATGGGGTGAGAGCCGACGTGCAT 544
 QY 61 ValIleSerThrProGlnAlaSerIlyAlaIleSerCysIlyGlnHisSerIleSer 80
 Db 545 GGGATTTCACCCCAACAGGCTTCACAGGAATCAGACAGAGCGCCAAACAGCATCTCA 604
 QY 81 TyrThrLeuSerArgAsnGlnThrValValIleGluTyrThrHisAspIlyAspThrAsp 100
 Db 605 TTCACATTTTCCAGAAAGTACAGACTGTGTGGTGGAATACACCATGACAGACACAGAT 664
 QY 101 MetPheGlnValIleArgSerThrGlnSerProIleAspPheValIleThrAspThrIle 120
 Db 665 ATGTTTCAGATTGGAAGGTCAACAGAAAGTCCCATTTGATTTGTAAGTACACAGGT 724
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 725 TCTGCACACAGAAAT--GATGAAACACAGATACACAAAGTACAAATCTTGATTTGG 781
 QY 141 CysArgIleValIleCysAspArgAsnGluProIlyThrThrAlaArgIlePheAlaIleGlyPhe 160
 Db 782 TCCAGATAGTCTGTGACAGAAACCTCGCTACACCGCTAGAAATATTTCTCGACGATT 841
 QY 161 AspSerSerIlyAsnIlePheLeuGlyGlyIlyAlaIleTyrIlyAsnProAspGly 180
 Db 842 GACTCTCCAAAAATATATTTCTTGGTGAGAAAGCGGCAAGTGGAAAGATCTGATGTGT 901
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleuValMetHisProArgIlyGlyPheThr 200
 Db 902 CATATGGAAGACATACTACCAATGAGTGTCTTGATGATCATCCGAAAGGAGGTTTACT 961
 QY 201 GluGlnSerGlnProGlyValIlyTrpArgIlyIleSerValIleCysGlyAspValIlyThrLeu 220
 Db 962 GAAAGATTCAGGCGCTGAGTATGCGCGGAAATTTCACTGTGAGATGTATATACATTA 1021
 QY 221 ArgGlyThrArgSerAlaGlnGlnArgGlyIlyIlyLeuValIleGlnSerGluThrAsnValLeu 240
 Db 1022 AGAGAAACCAAGATCTGCTCAGCAAAAGGCAAACTTGGAATGAAACCAATATTTCTG 1081
 QY 241 GlnAspGlySerIleuLeuAspLeuCysGlyAlaIleThrLeuLeuThrArgThrAlaAspGly 260
 Db 1082 CAGGATGGTTCCTTGTTACTATATGTGTGTCACCCCTTTGGGGAACACAGATGTGT 1141
 QY 261 LeuPheHisThrProThrGlnIlyHisIleGluAlaLeuArgGlnGlnIlyLeuAlaIle 280
 Db 1142 TTGCTCCATACACCACTCAAAAACATTTGAAGCTTTAGACGTGAATTAATGCTGCC 1201
 QY 281 ArgProGlnCysArgProValIleuAsnThrLeuAlaPheProSerIleAsnArgIlyGln 300
 Db 1202 AGGCGCTCAGTGCCTTGTGATTAACACTTTAGCTTTTCCAGATATCAACCGCAAAAGAT 1261
 QY 301 ValIleGlnGlnIlyGlnProThrAlaIlyTyrLeuSerCysGlyHisIleValHisGlyTyrHis 320
 Db 1262 GTCGTAGAAAGAAACAGCCTTGCGGTACCTCACTGTGTGATGTGACGCGCTATCAT 1321
 QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
 Db 1322 AACTGGGGCATCGCAGTACACAGAAAGTCAATGAACCGAGTGTCCATGTGTAGAAAC 1381
 QY 341 ValIleGlyProTyrValProLeuThrLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
 Db 1382 GTTGTGCTCTATATGTGCTCTATGTGCTTGTGTGAGAGAGATTTCACTAGATGCTGGA 1441

QY 361 ProProThrHisAlaPheThrProCysGlyHisIleValCysSerGlnIlySerAlaIlySTyr 380
 Db 1442 CCTCTACTACAGCTTTTAAACCTTTGTGACACGTCACTAGAAAATCTGCCAAATAT 1501
 QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIleCysProPheCys 400
 Db 1502 TGGTCTCAGATTTCACATGTCACATGTGTACACACCGGTTTACGTGCTTGTCTTCTGC 1561
 QY 401 AlaThrGlnLeuValGlyGlnGlnAsnCysAlleIlyLeuIlePheGlnGlyProIleAsp 420
 Db 1562 GCAATACAAATTAAGTACGTGAGAAAGAGCTGTGTCAAACTTAATTTCCAGGCGCTGTGAC 1621
 RESULT 12
 LOCUS BD190327 1717 bp DNA linear PAT 17-JUL-2003
 DEFINITION Elki phosphorylation related gene.
 ACCESSION BD190327
 VERSION BD190327.1 GI:3300066
 KEYWORDS MO 03008589-A/67.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1717)
 AUTHORS Matsuzaki, O., Matsuda, A., Nagano, Y. and Suzuki, N.
 TITLE Elki phosphorylation related gene
 JOURNAL Patent: WO 03008589-A 67 30-JAN-2003;
 ASHAI KASBI CORP,OSAMU MATSUZAKI,AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI
 COMMENT OS Mus musculus (mouse)
 PN WO 03008589-A/67
 PD 30-JAN-2003
 PF 15-JUL-2002 WO 2002JP007174
 PR 18-JUL-2001 JP 01P 218204,31-AUG-2001 JP 01P 263450 PR
 PI 21-JAN-2002 JP 02P 012176
 PI OSAMU MATSUZAKI,AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI PC
 C12N15/54,C12N9/12,G01N33/50,G01N33/15,C07K16/40,A61K39/385, PC
 A61K31/711
 CC Elki phosphorylation related gene
 FH key location/Qualifiers
 FT CDS location/Qualifiers
 (218)..(1474).
 FEATURES
 source 1..1717
 location/Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,55e-162 Length: 1717
 Score: 2111.50 Matches: 388
 Percent Similarity: 95.24% Conservative: 12
 Best Local Similarity: 92.38% Mismatches: 19
 Query Match: 92.21% Indels: 1
 DB: 6 Gaps: 1
 US-10-041-030-4 (1-420) x BD190327 (1-1717)
 QY 1 MetPheSerProGlyGlnGlnGlnIleTCysAlaProAlenlySGlnProValIleSTyrGly 20
 Db 1442 CCTCTACTACAGCTTTTAAACCTTTGTGACACGTCACTAGAAAATCTGCCAAATAT 1501
 QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIleCysProPheCys 400
 Db 1502 TGGTCTCAGATTTCACATGTCACATGTGTACACACCGGTTTACGTGCTTGTCTTCTGC 1561
 QY 401 AlaThrGlnLeuValGlyGlnGlnAsnCysAlleIlyLeuIlePheGlnGlyProIleAsp 420
 Db 1562 GCAATACAAATTAAGTACGTGAGAAAGAGCTGTGTCAAACTTAATTTCCAGGCGCTGTGAC 1621
 QY 41 SerArgPheAlaLeuTyrIleArgProIlyAlaAsnGlyValIleAspSerThrValHis 60
 Db 338 ACCAGATTTCCTTATTAAGGACCTTACCGCAAGTGTGTCAACCAAGCAATCCAC 397
 QY 61 ValIleSerThrProGlnAlaSerIlyAlaIleSerCysIlyGlnHisSerIleSer 80
 Db 1382 GTTGTGCTCTATATGTGCTCTATGTGCTTGTGTGAGAGAGATTTCACTAGATGCTGGA 1441

```

Db      398 ATGCTCCACACCAAGCGCTCCAGCCATCAGTCCAGAGACATCAGACATATCG 457
Qy      81  TTTTThLeuSerArgAngInThrValValGluTyrThrHisAspIysAspThrAsp 100
Db      458 TACACCTTTCACGAGCGACGAGACGAGTGTGAGTATACACACATTAAGACACAGAC 517
Qy      101 MetPheGlnValGlyArgSerThrgIuSerProIleAspPheValValThrAspThrIle 120
Db      518 ATGTTTCAAGTGGCGAGTCAACAGAAAGCCCATTTGACTTCTGTGTACAGACAGCGGT 577
Qy      121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrgIuSerThrIleSerArgPheAla 140
Db      578 TCCGGGGGTGACAGAACAGAT---GCCAGATCACACAGACCACTCTAGGTTCCGA 634
Qy      141 CysArgGleValCysAspArgAngInIuProTyrThrAlaArgIlePheAlaGlyPhe 160
Db      635 TCAGAGATCGTGTGTGACAGAAAGAGCATACACAGACGATATTCGCGGACGAGATTC 694
Qy      161 AspSerSerIysAsnIlePheLeuGlyGlyValAlaAlaIleThrgIuSerProAspGly 180
Db      695 GATCTTCCAAAATAATCTTTCTTGAGAGAAAGACGAAATGMAAAACCTGTATGGA 754
Qy      181 HisMetAspGlyLeuThrThraAngIyValIleValMetHisProArgGlyGlyPheThr 200
Db      755 CACATGATGAGACTCATCAACAGGTCTCTGTGTATCATCCCGGACAGAGAGCTTCACC 814
Qy      201 GluGluSerGlnProGlyValIleThraArgGluIleSerValCysGlyAspValIleThrLeu 220
Db      815 GAGGATATCCACAGCTGTGAGTGTGAGAGAGATCTCTGTGTGAGGATGTATACACTTG 874
Qy      221 ArgGluThraArgSerAlaGlnIuArgGlyValIleValGluSerGluThraAsnValIleu 240
Db      875 CAGAGACACAGGTGGCGCCAGACAGAGGAGAAAGCTGTGTGAAAGTGAACCAACGCTCG 934
Qy      241 GlnAspGlySerLeuIleAspLeuCysGlyValIleThrLeuThrTrpArgThrAlaAspGly 260
Db      935 CAAAGACGCTCTCTATTATCCTGTGTGGGGCCACTCTCTCTGGAGAACCGCAGATGGC 994
Qy      261 LeuPheHisThrProThrgIuIleValHisIleGluAlaLeuArgGlnIleIleAsnAlaAla 280
Db      995 CTTTTCACAGCTCTCTATCAGAAAGCACAATAGAAAGCCCTCGGACGAGATCAATGACGCC 1054
Qy      281 ArgProGlnIuSerProValGlyIleuAsnThrIleAlaPheProSerIleAsnArgIleVal 300
Db      1055 CAAACCCAGATGCGCGTGGGCTTAAACACCTGGGCTTCCCAAGCATCAACCGGAGGAA 1114
Qy      301 ValValGluGluIleGlnProTrpAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
Db      1115 GTGGTGAAGAGAAAGACGCCCTGGGATACCTGAGCTGGCGCATGTGCACGGCTATCCAC 1174
Qy      321 AsnTrpGlyHisIleArgSerAspThrgIuAlaAsnGluArgGluCysProMetCysArgThr 340
Db      1175 CACTGGGGGCATCGAGGACAGACAGAAAGCAAGAGAGGAGTGTCCATGTGATGACT 1234
Qy      341 ValGlyProTyrValProleuThrPheGlyCysGlyAlaGlyPheTyrValAspAlaGly 360
Db      1235 GTGGGCGCTTACGCTCTCTGTGGCTGTGAGGCGAGATTATGTATGCAANNNNNN 1294
Qy      361 ProPheThrHisAlaPheThrProCysGlyHisValCysSerGluIleSerAlaIleVal 380
Db      1295 NNNNNNNNCTCATGTTTTCACCCCTGGGCGACGTGTCTTCAGAGAAAGCTTCCCAAGTAC 1354
Qy      381 TrpSerGlnIleProleuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db      1355 TGGTGGCAGATCCCACTGCCCAAGGACGACGAGCGCTTCAAGCCCTGTGCTGCTCAG 1414
Qy      401 AlaThrGlnLeuValGlyGluGlnAsnGlyIleValLeuIlePheGlnIleProIleAsp 420
Db      1415 GCCAGCGACCTGTGTGTGAAAGAACTCAATTAATTAATTTTCCAAAGGTCAGGTGAC 1474

```

RESULT 13
AF302504
LOCUS

AF302504 1734 bp mRNA linear ROD 05-JUL-2001

```

DEFINITION Mus musculus pellino 2 (Peli2) mRNA, complete cds.
ACCESSION AF302504
VERSION AF302504.1 GI:10242356
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
REFERENCE
AUTHORS Resch, K., Jockusch, H. and Schmitt-John, T.
TITLE Assignment of homologous genes, Peli1/Peli1 and Peli2/Peli2, for the Pelle adaptor protein Pellino to mouse chromosomes 11 and 14 and human chromosomes 2p13.3 and 14q21, respectively, by physical and radiation hybrid mapping
JOURNAL Cytogenet. Cell Genet. 92 (1-2), 172-174 (2001)
MEDLINE 21203570
PubMed 11306823
REFERENCE 2 (bases 1 to 1734)
AUTHORS Resch, K., Jockusch, H. and Schmitt-John, T.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-2000) Developmental Biology and Molecular Pathology, University of Bielefeld, Universitaetsstrasse 25, Bielefeld, NRW 33615, Germany
FEATURES
source location/Qualifiers
1..1734
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
1..1734
/gene="Peli2"
218..1477
/gene="Peli2"
/note="similar to Drosophila melanogaster Pellino"
/codon_start=1
/product="pellino 2"
/protein_id="AAG15392.1"
/db_xref="GI:10242357"
/translation="MPSFGQEPSPAPNPKPVKRELVVLGYNGALPNDGRGKKSRFTLYKRYASGVKPSSTIHMTSPQASKALISRGHNSIYLSRSQTVVVEYTHDKDPMQVGRSTESPIDFVDTVSGGONEDAOITOSTISRPACRIYCDRNEPYARIPLAAGFDSSKNTFLGKAKAKMNPQGHMDGLTNCVILVNHPEGFTSESQGVMEIRISVGIVYTLEKTSADQQRGLVSEFNVDGSLIDLCATLLMRPADGIFAPYTOCHITALEOETNAARPOQCVGLNTTLAFPSINREVEEKQPMAYLSCGVHGHGHNHSDDTANERECFMCMTGVPEVPLMLGCEAGFYVXXXXXHVFPCHVCEKSKAKYMSQIPLPHGTHAFHNAECFATQLVGEQNCIKLIFGQVFD"
ORIGIN
Alignment Scores:
Pred. No.: 1,566-162 Length: 1734
Score: 2111.50 Matches: 368
Percent Similarity: 95.24% Conservative: 12
Best Local Similarity: 92.38% Mismatches: 19
Query Match: 92.21% Indels: 1
Gaps: 1
DB:
US-10-041-030-4 (1-420) x AF302504 (1-1734)
Qy      1 MetPheSerProGluGlnGluIleIscYsAlaProAsnIleGluProValIleTyrGly 20
Db      218 ATGTTTTCCTCCGGGCGACAGAGAACCCAGCCGCCCAACAGAGACCGGTGAATAACAG 277
Qy      21 GluLeuValIleGluIleTyrAsnGlyValIleuProAsnGlyAspArgGlyArgArgIle 40
Db      278 GAGCTGTGTGCTCTGGGGAACAATGTCTTAACTTAATGGTGAACAGGGGCAAGAA 337
Qy      41 SerArgPheAlaLeuTyrIleArgProIyAlaAsnGlyValIleProSerThrValHis 60
Db      338 AGCAATTTTACCTTATTAAGCGGACCTACGCAATGTGTCTAAACCAAGCAATATCCAC 397
Qy      61 ValIleSerThrProGlnAlaSerIysAlaIleSerCysIleGlyGlnIleSerIleSer 80
Db      398 ATGCTCCACACCAAGCGCTCCAGGCCATCAGCTCCAGAGACATCAGACATATTCG 457

```

```

QY      81  TTTTThleuSerArgAsnGlnThrValValGluTyrThrHisAspLysAspThrAsp 100
Db      458  TACACGTTGTCCAGGACCGAGCGGTAGTGGTGGAGTACACACAGATTAACACAGC 517
QY      101  MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
Db      518  ATGTTTCAGGTGGGAGGTCAACAGAAAGCCCATTTGACTTGCTGTGTCACACACGGTT 577
QY      121  SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      578  TCCCGCGGTGACAGAACGAAGAT---CCCGAGATCACACAGAGACCATCTCTGTGTCGA 634
QY      141  CysArgGlnIleValCysAspArgAsnGlnLysProTyrThrAlaArgIlePheAlaIleGlyPhe 160
Db      635  TGCAGAGATGTTGTGACAGGACGACGATACACAGCAGCATATTGCGCGGAGAGATTC 694
QY      161  AspSerSerLysAsnIlePheLeuGlyGluLysAlaAlaLysTrpLysAsnProAspGly 180
Db      695  GATTCTTCCAAAATAATATCTTTCTTGGAGAGAAAGCAGCAAAATGGAATAACCTGATGA 754
QY      181  HisMetAspGlyLysLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db      755  CACATGATGACTCACTACCAACCGTGTCTGTGATGTCACCCGCAAGAGAGCTTCAACC 814
QY      201  GlnGluSerGlnProGlyValLysTrpArgGlnIleSerValCysGlyAspValTyrThrLeu 220
Db      815  GAGGATATCCAGCCCTGGAGTCTGGAGAGAGATCTGCTGTGGGAGTGTATACCTTGG 874
QY      221  ArgGlnThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGlnLysThrAsnValLeu 240
Db      875  CGAGAGACCAAGTCCGGCCAGCAGAGAGGGAAGCTGTGGAAAGTGAAGCACAACCTCTG 934
QY      241  GlnAspGlySerLeuIleAspLeuCysGlyValAlaThrLeuLeuTrpArgThrAlaAspGly 260
Db      935  CAAGCGGCTCCCTCATTTGACTGTGTGGGCGACTCTCTCTGTGAGAAACCCCAATGGCC 994
QY      261  LeuPheHisThrProThrGlnLysHisIleGlnAlaLeuArgGlnGlnIleAsnAlaIle 280
Db      995  CTTTTCACGCTCTCACTCAGAAAGCATAGAAAGCCCTCCGCGAGAGATCAATGACAGCC 1054
QY      281  ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGln 300
Db      1055  CGAACCCAGTGGCCCGTGGGCTTAAACACCTGGGCTTCCAGCATCAACCGGAAGGAA 1114
QY      301  ValValGlnGluLysGlnProThrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
Db      1115  GTGGTGAAGAGAACAGCCCTGGGCTATACCTGAGCTGGGCGCATGTGCACGCTAACAC 1174
QY      321  AsnTrpGlyHisArgSerAspThrGlnAlaAsnGlnLysArgLysCysProMetCysArgThr 340
Db      1175  CACTGGGGCGCATCGAGCGACACAGAACGCAACGAGAGGAGTGTCCCATGTGCTAGACT 1234
QY      341  ValGlyProTyrValProLeuTrpLeuGlyCysGlnAlaGlyPheTyrValAspAlaGly 360
Db      1235  GTGGGCGCCCTAGCTCTCTGTGGCTGTGAGCGAGATTTTATGTGANNNNNNN 1294
QY      361  ProProThrHisAlaPheThrProCysGlyHisValCysSerGlnLysSerAlaLysTyr 380
Db      1295  NNNNNNNCTCAGTGTTCACCCCTGGCGGCGCTGTGTCGAAAGCTGCCAAGAGAC 1354
QY      381  TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db      1355  TGTGTGCGAGATCCCACTGCCCAAGGAGCAGCGCTTTCACGCGCTGTCCGTTCTGCC 1414
QY      401  AlaThrGlnLeuValGlyGlnGlnAsnGlyIleLysLeuIlePheGlnGlyProIleAsp 420
Db      1415  GCCACGCGAGCTGTGTGTGAACAGAACTGCATCAATTGATTTTCCAAAGTCCAGTGGAC 1474

```

```

RESULT 14
BC075973      BC075973      2474 bp      mRNA      linear      VRT 27-JUL-2004
LOCUS        Dantio rerio zgc:92268, mRNA (cDNA clone MGC:92268 IMAGE:7053197),
DEFINITION

```

```

ACCESSION      complete cds.
VERSION        BC075973
KEYWORDS      MGC.
SOURCE        Dantio rerio (zebrafish)
ORGANISM      Dantio rerio
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
                Cypriniformes; Cyprinidae; Danio.
                1 (bases 1 to 2474)
                Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
                Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
                Altschuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
                Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
                Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
                Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
                Scheetz, T.E., Brownstein, M.J., Uedl, T.B., Toshiyuki, S.,
                Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
                Abramson, R.D., Malek, J.A., Gunaratne, P.H., Richards, S.,
                McKernan, K.J., Males, J.A., Gunaratne, P.H., Richards, S.,
                Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultk, S.W.,
                Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
                Fahey, J., Hellon, E., Kettelman, M., Madan, A., Rodriguez, S.,
                Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
                Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
                Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
                Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
                Schermer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
                Generation and initial analysis of more than 15,000 full-length
                human and mouse cDNA sequences
                Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
                1247932
                2 (bases 1 to 2474)
                Strausberg, R.
                Direct Submission
                Submitted (06-JUL-2004) National Institutes of Health, Mammalian
                Gene Collection (MGC), Cancer Genomics Office, National Cancer
                Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
                USA
                NIH-MGC Project URL: http://mgc.nci.nih.gov
                Contact: MGC help desk
                Email: gcgabs-remail.nih.gov
                Tissue Procurement: Len Zon, Harvard
                cDNA Library Preparation: Open Biosystems
                DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Sequencing Group at the Stanford Human Genome
                Center, Stanford University School of Medicine, Stanford, CA 94305
                Web site: http://www.shgc.stanford.edu
                Contact: (Dickson, Mark) mcd@paxil.stanford.edu
                Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
                R. M.
                Clone distribution: MGC clone distribution information can be found
                through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
                Series: IPAK Plate: 179 Row: m Column: 7
                This clone was selected for full length sequencing because it
                passed the following selection criteria: Hexamer frequency ORF
                analysis, similarity but not identity to protein.
                Location/Qualifiers
                1..2474
                /organism="Dantio rerio"
                /mol_type="mRNA"
                /db_xref="taxon:7955"
                /clone="MGC:92268 IMAGE:7053197"
                /tissue_type="whole, adult zebrafish"
                /clone_lib="NIH_ZGC_7"
                /lab_host="DH10B"
                /note="vector: pExpress1"
                1..2474
                /gene="zgc:92268"
                /db_xref="locusID:436899"
                218..1504
                /gene="zgc:92268"

```

```

/codon_start=1
/product="zgc:92268"
/protein_id="AAH75973.1"
/db_xref="GI:49903831"
/translation="MFSPOEHCAPTKEPVKKYKGLVVLGVNGLPNGRGRKSRFA
LYKRAANGVKSSTVHILNTPKASKAVNKKQHSISYTSRNTQVTVVEVSHDTPMF
QGRSTREPIIDVYVDTVTVSGGSEETPTQSTISRPACVCEENPPTTARIIPAAGF
DSSKNIPLFEKAKAKMNPDPGHMDGLTTNGVLVHMPRGFTPEESKPGVMEISVCGDVY
TLRETSAGTRGLVSESNVLQDGLVLDLCAATLMRTADGLFHTPTQKHLALOE
LNAARQCEVGLNTLAFPSMORCSRALSNVPAQEDQPMVYLACGHVGHMDGHSRE
RESNAORECPMCRVAGPVYPLMVGCEPAFYVDGAPTHFVPCGHVCSKSRVYAEI
PLPHGTHAHCPCFATQTLNLTQKAKLIFQSPID"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 9,59e-154 Length: 2474
Score: 2006.00 Matches: 361
Percent Similarity: 90.89% Conservative: 28
Best Local Similarity: 84.35% Mismatches: 31
Query Match: 87.60% Indels: 8
DB: 5 Gaps: 2

```

US-10-041-030-4 (1-420) x BC075973 (1-2474)

```

QY 1 MetPheSerProGlyGlnGluGluHsCybAlaProAsnLysGluProValLysTyrgly 20
DB 218 ATGTTTTCACCGGGTACAGAGAGACCTGTGCGCCCAAGAGCCAGTAAAGTACGGG 277
QY 21 GluLeuValValLeuGlyTyraGnglyAlaLeuProAsnGlyAspArgGlyAspGly 40
DB 278 GAACCTGCTGTTTAAAGATACATGCTCTCCCAATGCTGACAGAGGCCAAGAA 337
QY 41 SerArgPheAlaLeuTyrylAspArgProLysAlaAsnGlyValLysProSerThrValHs 60
DB 338 AGCCGGTTTGCCCTTTACAAAGAGACCAAGCCAAATGGGTCAGGCCAGCACTGAC 397
QY 61 ValLysSerThrProGlnAlaSerLysValAlaLysSerCysLysGlyGlnHsSerIleSer 80
DB 398 ATCTTTCACACACACACACACACACACACACACACACACACACACACACACAC 457
QY 81 TyThrLeuSerArgAsnGlnThrValValValGluTyThrHsAspLysAspThrAsp 100
DB 458 TACACGCTTCCAGAAACACAGAGGTGTGTGAGTACAGCATACAAAGACAGAC 517
QY 101 MetPheGlnValGlyArgSerThrGlySerProLysAspPheValValThrAspThrIle 120
DB 518 ATGTTCCAGATCGCGCATCGACGAGGGCCGATTGACTTTGTGTGACTGACACTG 577
QY 121 SerGlySerGlnAsnThrAspGlyValAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 578 TCAGGGGAGAGAGAGTGAAGAGACACCCATCATCTCAAGCACATCTCCCGCTCGCC 637
QY 141 CysArgIleValCysAspArgAsnGluProTyThrAlaArgIlePheAlaIleGlyPhe 160
DB 638 TGCCGAGTGTGTGTGAGGTAAACCACTTACACCGCACGAAATCTTCGTGAGGCTTC 697
QY 161 AspSerSerLysAsnIlePheLeuGlyGlyLysAlaAlaLysTyThrLysAsnProAspGly 180
DB 698 GATTTCTTCAAGAAATATATTTCTCGAGAAAAAGACCCAAATGAGAAATCAGATGT 757
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
DB 758 CACATGAGCGTTTGACTACATGAGGTGTGTGATCATCACCGCGGGCGGTTCACA 817
QY 201 GluGluSerGlnProGlyValTyraArgLysLeuSerValCysGlyAspValTyThrLeu 220
DB 818 GAGGAATCCAGCTGTGTGTGTGAGAGAGATTCAGTGTGTGAGATGTTTACACACTG 877
QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
DB 878 AAGAAACACGCTCTCAAACTCGAGAAAAATGTGTGAGAGTGAAGTAAAGTTCGTG 937

```

```

QY 241 GlnAspGlySerLeuIleAsnLeuCybGlyAlaThrLeuLeuThrPheArgThrAlaAspGly 260
DB 938 CAGGAGGCGCTCTGTGTGATCACTTTGTGAGAAACCTCTCTGGGCGACCGCTGACGGC 997
QY 261 LeuPheHsThrProThrGlnLysHsIleGlyAlaLeuAspGlnGluIleAsnAlaIle 280
DB 998 CTTTTCACACCCCCACAGAAAGACCTGAGAGCCCTTCGGCAGAGCTGAACGACACC 1057
QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArg----- 298
DB 1058 CGGCCCCAGTGCCTGTGGCCCTCAACACACTGGCCCTTCCAGCATGACAGCTGCAGT 1117
QY 299 -----LysGluValValGluGluLysGlnProThrAlaTyThrLeuSerCys 313
DB 1118 CGTGCCCTGAGCAACGCTCCCTGCCAGAAACAAACACCCCTGGGTCTACCTGCTGT 1177
QY 314 GlyHisValHisGlyTyThrHsAsnThrPheLysArgSerAspThrGlnAlaAsn---Glu 332
DB 1178 GGTGATGTTTCAATGCTACACACAGCTGGGCCACCATCAGAGGGAGTCCAAATGCCCA 1237
QY 333 ArgGluCysProMetCysArgThrValGlyProTyThrValProLeuThrLeuGlyCysGlu 352
DB 1238 AGAGAGTGCCTCATGTGAGAGGCGCGGTCCCTATGTCCACTGTGCTGGCTGGATGTGA 1297
QY 353 AlaGlyPheTyThrValAspAlaGlyProProThrHsAlaPheThrProCysGlyHisVal 372
DB 1298 CCAGCTTTCAGTGAACACTGTGTGGCCACACACAGTCTTGATCTGTGTGCGCATGTG 1357
QY 373 CysSerGluLysSerAlaLysTyThrPheSerGlnIleProLeuProHisGlyThrHsAla 392
DB 1358 TGCTCGAAAGATCCACCGGCTATTGGCAGAGATCTCTTGGCAGATGACACCGCCGCT 1417
QY 393 PheHisAlaAlaCysPheProPheCysAlaThrGlnLeuValGlyGluGlnAsnCysIleLys 412
DB 1418 TTCACAGCTGCTGTCTTGTGTGCTACTCACTCAACCTGACACAAAGTGGCCAAAG 1477
QY 413 LeuIlePheGlnGlyProIleAsp 420
DB 1478 CTCATCTTTCAGGGGCTTATGAC 1501

```

RESULT 15

AR483566

LOCUS AR483566 1257 bp DNA linear PAT 14-MAY-2004

DEFINITION Sequence 3 from patent US 6703487.

ACCESSION AR483566

VERSION AR483566.1 GI:47246650

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1257)

AUTHORS Bird,T.A. and Cosman,D.J.

TITLE Humanellino polypeptides

JOURNAL Patent: us 6703487-A 3 09-MAR-2004;

FEATURES

location/Qualifiers

1..1257

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 7,63e-147 Length: 1257

Score: 1917.00 Matches: 342

Percent Similarity: 90.48% Conservative: 38

Best Local Similarity: 81.43% Mismatches: 38

Query Match: 83.71% Indels: 2

DB: 6 Gaps: 1

US-10-041-030-4 (1-420) x AR483566 (1-1257)

QY 1 MetPheSerProGlyGlnGluGluHsCybAlaProAsnLysGluProValLysTyrgly 20

DB 1 ATGTTTTCCTGATCAAGAAATATCAT-----CATCTAAAGCACACGATAAATATGTGT 54

QY 21 GIuLeuValIValLeuGIYThraSngIYAlaLeuProaSngIYAspArgIYArghArgIys 40
 Db 55 GAACTCATGTGCTTGGATATATAGATCTCCAAACGGATATAGAGAAAGAGAA 114
 QY 41 SerArgPheAlaLeuTYrLYsArgProLYsAlaAsngIYValLYsProSerThraValHis 60
 Db 115 AGTAGGTTGGCTTTGTTTAAAGACTTAAGCAATGGGGTGAAGCCGACGACTGGCAT 174
 QY 61 ValIleSerThProGlnAlaSerLYsAlaIleSerCYsLYsGlyGlnHisSerIleSer 80
 Db 175 ATYTGCTTGACTCTCAGGCTGCAAGCAATAGCAACAAAGACGAGATGACATATCA 234
 QY 81 TYrThrLeuSerArgAsnGlnThrValValIGluTYrThiAspLYsAspThrAsp 100
 Db 235 TATACTTATCTCGGCCCAAGACTGTGGTGGTGAATATCTCATGACAGCAACCCGAT 294
 QY 101 MetPheGlnValIGLYsSerThrgIuSerProIleAspPheValValThraSPThrIle 120
 Db 295 ATGTTTCAGATTGGCCGGTCGACTGAAGCCCATGATTTTGTAGTACTGACAGGTT 354
 QY 121 SerGlySerGlnAsnThraSPgluAlaGlnIleThrgIuSerThrIleSerArgPheAla 140
 Db 355 CCTGGAAGCAAGTAATTCTGATACACAGTCAGTCAAGCACTATATCAAGATTTCGC 414
 QY 141 CYsArgIleValCYsAspArgAsnGluProTYrThraAlaArgIlePheAlaIAGLYPhe 160
 Db 415 TGCAATATCATATGGAAGGAAATCTCCCTTACAGCACGAAATTTATCTGACAGGTTT 474
 QY 161 AspSerSerLYsAsnIlePheLeuGIYgluLYsAlaIalysTrpLYsAsnProAspGly 180
 Db 475 GACTCATCAAAATACTCTTCTTGGGGGAAGGCTGCCAAATGGAAATGAAATCATGATGA 534
 QY 181 HisMetAspGlyLeuThrThraSngIYAlaLeuValMetHisProArgGlyGlyPheThr 200
 Db 535 CAGATGGATGGCTTGACCACTAATGGTGTCTGTGATGATCCACGCAATGGGTTCA 594
 QY 201 GIuGluSerGlnProGlyValTrpArgGluIleSerValCYsGlyAspValTYrThrLeu 220
 Db 595 GAAAGACTCCAAAGCTCGAATATGAGAGAAATATCGGTGTGGGAATGATTTAGCTTA 654
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLYsLeuValGluSerGluThraSngValLeu 240
 Db 655 CGTGAATCCAGATCGGCTCAGCAGAGAGAAATGTGTGAATGGAATCAATCAGTTA 714
 QY 241 GlnAspGlySerLeuIleAspLeuCYsGlyAlaThrLeuLeuTrpArgThraAlaAspGly 260
 Db 715 CAAGATGGCTGTATTAATGACCTGTGTGCAACATTTGTATGGGTACTGCAGAAAGGC 774
 QY 261 LeuPheHisThrProThrgIuLYsIleGlnAlaLeuArgGlnGluIleAsnAlaAla 280
 Db 775 CTTTCCCAACACTCCACCGTGAAGCATTTAGAAGCTTTAAGACAGAAATCAATGCAGCA 834
 QY 281 ArgProGlnCYsProValIGLYsLeuAsnThrLeuAlaPheProSerIleAsnArgLYsGlu 300
 Db 835 CGACCTCAGTGCCCTGTAGGGTTCAACACATGACATTTCTAGTATGAAGAGAAAGAC 894
 QY 301 ValValIGluGluLYsGlnProTrpAlaTYrLeuSerCYsGlyHisValHisGlyTYrHis 320
 Db 895 GTTGATGATGAAGAACAAACATGATATCTAAACTGCGGCATGTACATGACTATCAT 954
 QY 321 AsnTrpGlyHisArgSerAspThrgIuAlaAsngIYAlaArgGluCYsProMetCYsArgThr 340
 Db 955 AACTGGGGAACAAGAAAGAACAGTATGGCAAAAGATCGTAATGTCCTATGTAGTCT 1014
 QY 341 ValGlyProTYrValProLeuTrpLeuGIYCYsGluAlaGlyPheTYrValAspAlaGly 360
 Db 1015 GTTGGTCCCTATGTCTCTGTGGCTTGATGAAAGTGAATTTATGTGACGCGGCG 1074
 QY 361 ProProThrHisAlaPheThrProCYsGlyHisValCYsSerGluLYsSerAlaLYsTYr 380
 Db 1075 CCTCAACCAATGCTTAGCCGCTGGGCAATGTGTTCAGAAAAAGCAACTGCTAT 1134

QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCYsAspPheCYs 400
 Db 1135 TGGTCCAGATCCCACTTCTCATGTACTACTTTTCAATGACGCTGTCTTTGT 1194
 QY 401 AlaThrGlnLeuValIGLYGluGlnIleAsnCYsIleLYsLeuIlePheGlnGlyProIleAsp 420
 Db 1195 GCACATCAGTTGGCTGTGGAACAAAGCTACATCAGACTTATTTTCAAGAACCTCTAGAC 1254

Search completed: December 9, 2004, 15:53:39
 Job time : 7563 secs

| | | | |
|----|------|---|------|
| Db | 61 | GAGCTGGTGGTCTCGGGTACAAATGGTGGCTTTACCAAGAGATGAGAGCGAGAA | 120 |
| Qy | 41 | SeATgPheAlaLeuTYrLYsArGPoLYsAlaAsnGlyValLYsProSerThrValHis | 60 |
| Db | 121 | AGTAGATTTGGCCCTCTACAAAGGGGCCAAGGAAAGGTGTCAAAACCGACCGCTCCAT | 180 |
| Qy | 61 | ValLYsSerThrProGlnAlaSerLYsAlaLYsSerCYalysGlyLYsGlnLYsSerLYsSer | 80 |
| Db | 181 | GTGAATATCCACCCTCCAGGCGATCCAAAGGCTATTCAGCTGTCAAGAGTCAACAGTATATCC | 240 |
| Qy | 81 | TYrThrLeuSerPheArgAsnGlnThrValValValGluTYrThrHisAspLYsAspThrAsp | 100 |
| Db | 241 | TACACTTGTCTCAAGGAATCAGACTGTGTGGTGGAGATCAACATGATAAAGTTACGGAT | 300 |
| Qy | 101 | MetPheGlnValGlyArgSerThrGluSerProLYsAspPheValValThrAspThrLe | 120 |
| Db | 301 | ANGTTTCAGTGTGGGCGAGATCAACAGAAACCCCTATGCACTTGGTGTCTCACAGACAGATT | 360 |
| Qy | 121 | SerGlySerGlnAsnThrAspGlnValGlnLYsIleThrGlnSerThrLYsSerArgPheAla | 140 |
| Db | 361 | TCGGCAGCCAAAGAACCGAGACAGAGCCCGATCACACAGACCATATCCAGATTCCGC | 420 |
| Qy | 141 | CysArgLYsValCysAspArgAsnGluProTYrThrAlaArgLYsPheAlaAlaGlyPhe | 160 |
| Db | 421 | TGCAGAGATCGTGTGGCAGAGATGAACCTTACACAGACGGAATATTCCTCCGCGGATTT | 480 |
| Qy | 161 | AspSerSerLYsAsnLYsPheLeuGlyGlyLYsValAlaLYsThrLYsAsnProAspGly | 180 |
| Db | 481 | GACTCTTCCAAAACATATTTCTTGGAGTAAAGGACGCAATGTGAAAACCCCGACGGC | 540 |
| Qy | 181 | HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgLYsGlyPheThr | 200 |
| Db | 541 | CACATGGATGGGCTCACTACTATGGCGTCTGTGGTATGCATCCACAGAGGGGCTTCACC | 600 |
| Qy | 201 | GluGluSerGlnProGlyValLYsThrArgGluLYsSerValCysGlyAspValTYrThrLeu | 220 |
| Db | 601 | GAGGAGTCCACAGCCCGGGGCTGTGGCGCGAGATCTGTCTGTAGAGATGTGATACACTTG | 660 |
| Qy | 221 | ArgGluThrArgSerAlaGlnGlnArgGlyLYsValSerValGluSerGluThrAsnValLeu | 240 |
| Db | 661 | CGAGAAACCAAGGTGGGCCCGACGACGAGGAAGAGCTGTGAAGTGAAGCAACGCTCTG | 720 |
| Qy | 241 | GlnAspGlySerLeuLYsAspLeuCYsGlyValAThrLeuLeuThrArgThrAlaAspGly | 260 |
| Db | 721 | CAGGACGGGTCTCTCATTTGACTGTGTGGGCGCATCTCTCTGAGAAACAGCAGATGGG | 780 |
| Qy | 261 | LeuPheHisThrProThrGlnLYsHisLYsGlnValAlaLeuArgGlnGluLYsAsnAlaAla | 280 |
| Db | 781 | CTTTTTCATACCTCCAACTCAGAAAGCAATAGAAAGCCCTCGGACAGGAGATTAACCGCGCC | 840 |
| Qy | 281 | ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerLYsAsnArgLYsGlu | 300 |
| Db | 841 | CGGCTCATGTGCTCTGTGGGGCTCAACACCTGGGCTTCTCCAGCATCAACAGGAAGAG | 900 |
| Qy | 301 | ValValGluGlnLYsGlnProThrAlaTYrLeuSerCYsGlyHisValHisGlyTYrHis | 320 |
| Db | 901 | GTGGTGGAGGAAAGACAGCCCTCGGGGATATCTCAAGTTGTGGCACTGACAGGGATACAC | 960 |
| Qy | 321 | AsnTYrGlyHisArgSerAspThrGlnAlaAsnGluArgGlyCysAspMetCysArgThr | 340 |
| Db | 961 | AACTGGGGCCATCGGAGTGACACGGAGGCCCAAGAGGGAGTGTCCTCATGTGACAGACT | 1020 |
| Qy | 341 | ValGlyProTYrValProLeuThrLeuGlyCysGlnValGlyPheTYrValAspAlaGly | 360 |
| Db | 1021 | GTGGGCCCCCTATGTGCTCTCTGTGGCTGTGTGAGGACGAGATTTATATGAACGACGGA | 1080 |
| Qy | 361 | ProProThrHisAlaPheThrProCYsGlyHisValCysSerGlyLYsSerAlaLYsTYr | 380 |
| Db | 1081 | CCGCCAACTCATGCTTTCATCTCCCTGTGGACACGCTGTCTCGGAGAAAGTCTGCAAAATTC | 1140 |
| Qy | 381 | TrpSerGlnLYsProLeuProHisGlyTYrHisValAspHisValAlaCysAspProPheCys | 400 |

| | | | |
|---|------|---|------|
| Db | 1141 | TGSHCTCAAGATCCCGTTGCTGCTCAATGCAATTCATTCATTTACAGCGTGCCTTTCTGT | 1200 |
| Qy | 401 | AlaThrGluLeuValGlyGluGlnAsnCysIleLeuLeuIlePheGlnGlyProIleAsp | 420 |
| Db | 1201 | GCTACACAGCGTGGTGGGAGCAAAATCGATCAATTAATTTTCCAAAGGTCGAATTGAC | 1260 |
| RESULT 2 | | | |
| US-09-843-905A-5 | | | |
| ; Sequence 5, Application US/09843905A | | | |
| ; Patent No. 6703487 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Bird, Timothy A. | | | |
| ; APPLICANT: Cosman, David J. | | | |
| ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES | | | |
| ; FILE REFERENCE: 2390-A | | | |
| ; CURRENT APPLICATION NUMBER: US/09/843, 905A | | | |
| ; PRIORITY FILING DATE: 2001-04-27 | | | |
| ; PRIOR APPLICATION NUMBER: US 60/200,198 | | | |
| ; PRIOR FILING DATE: 2000-04-28 | | | |
| ; NUMBER OF SEQ ID NOS: 15 | | | |
| ; SOFTWARE: PatentIn version 3.1 | | | |
| ; SEQ ID NO 5 | | | |
| ; LENGTH: 1260 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Mus musculus | | | |
| US-09-843-905A-5 | | | |
| Alignment Scores: | | | |
| Pred. No.: 2,066-231 | | | |
| Score: 2175.50 | | | |
| Percent Similarity: 97.38% | | | |
| Best Local Similarity: 95.00% | | | |
| Query Match: 95.00% | | | |
| DB: 4 | | | |
| Gaps: 1 | | | |
| US-10-041-030-4 (1-420) x US-09-843-905A-5 (1-1260) | | | |
| Qy | 1 | MetPheSerProGlyGlnGluGlnIleCysAlaProAsnIleGluProValIleGlyTrpGly | 20 |
| Db | 1 | ATGTTTTCCTCCGGGCCAGAGAAACCCAGGCCCCCAACAGAGCCGGTGAATATACGGG | 60 |
| Qy | 21 | GluLeuValIleLeuGlyTrpAsnGlyAlaLeuProAsnGlyAspArgGlyArgGlyAs | 40 |
| Db | 61 | GAGCTGGTGGCTCGGGGGTCAATGAGTGGCTTACCTAATGATGTGACAGGGGACAGAGAAA | 120 |
| Qy | 41 | SerTrpGheAlaLeuTrpIleValArgProIleAlaAsnGlyValIleProSerThrValHis | 60 |
| Db | 121 | AGCAGATTTCGCTCTTAAGCGGACCTACGCGAGTGTGTCAAAACCCAGACAAATCCAC | 180 |
| Qy | 61 | ValIleSerThrProGlnAlaSerIleValIleSerCysIleGlyGlnHisSerIleSer | 80 |
| Db | 181 | ATGGTCTCCACACACAGCGGCTCCAGGCCATCACTCCAGAGGACATCACAGCATATCG | 240 |
| Qy | 81 | TyrThrLeuSerArgAsnGlnThrValIleValGluTrpThrHisAspIleAspThrAsp | 100 |
| Db | 241 | TACACGTTGTGCACGAGACCCAGACCGTACGTAGTGGAGTAACACACAGATTAAGACACGAG | 300 |
| Qy | 101 | MetPheGlnIleGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle | 120 |
| Db | 301 | ATGTTTACGGTGGGCAAGTCAACAGAAAGCCCAATTCATCTTGGTGCACAGACGCGTT | 360 |
| Qy | 121 | SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla | 140 |
| Db | 361 | TCCGGCGGTACAGAACGAAT--GCCCAATCACACAGACACCATCTTTAGGTTCCGA | 417 |
| Qy | 141 | CysArgIleValCysAspArgAsnGluProTrpThrAlaArgIlePheAlaIleGlyPhe | 160 |
| Db | 418 | TGCAGGATCGTGTGTACAGGAACGACCATATACAGCAGCATATTCGGGGCAGATTC | 477 |
| Qy | 161 | AspSerSerIleValIlePheLeuGlyGluLeuAlaIleAlaIleTrpIleAspProAspGly | 180 |
| Db | 478 | GATCTTCCAAAAATATCTTTCTTGGAGAGAAAGACAGCAAAATGAGAAAAACCCGTAGTGA | 537 |


```

QY      181 HlMeArbglYleuthrThraenglyValleuValMeChlProArnglyglYpHeThr 200
      538 CACATGATGAGTCACTACCAATGATGCTCTGATGATGACCCGAGAGAGGCTTCACC 597
QY      201 GluGlusErGlnProGlyValTPArGgJuileSeValCyseGlyAspValTYrThrLeu 220
      598 GAGGAATCCAGAGCTGAGTCTGGAAGAGATCTCTGTGTGGAGATGTGTACACTTG 657
QY      221 ArgGluThrArgSerAlaGlnGlnArglyLyLeuValGluSerGluThraenValleu 240
      658 CGAGAACCCAGAGTCCGAGCCAGAGAGGAGAAAGCTGTGGAAGTGAACCAACGCTCG 717
QY      241 GlnAerGlySerLeuIleAspLeuCyseGlyAlaThrLeuLeuTPArGThrAlaAspGly 260
      718 CAGAGGCTCTCTCACTTGACTGTGTGGGCGCACTCTCTCGAGAACCGAGATGGC 777
QY      261 LeuPheHleThrProThrGlnlyVshIleGluAlaLeuArnglInglIleAenAlaIa 280
      778 CTTTTCAGCGCTCTTACTCAAGAGCAATGAGACCTCCGAGAGAGATCAATGACGCC 837
QY      281 ArgProGlnCyseProValGlyLeuAenThrLeuAlaPheProSerIleAenArglyGlu 300
      838 CGACCCCAAGTGGCCCGTGGGCTTTAAACACCTGGCTTCCAGCATCAACCGAAGAA 897
QY      301 ValValGluGluIlyseGlnProTPArIleTYrLeuSerCyseGlyVhIleValHleGlyTYrHis 320
      898 GTGTGGAAGAAGAGAGCGCTTGGCATACCTGAGCTGCGCATGTGCAGGCTACAC 957
QY      321 AenTPRGlyVhIleArgSerAenThrGluAlaAenGluArnglIlyCyseProMetCyseArgThr 340
      958 AGCTGGGGCCATCGAGACCAAGCGGAGAACCAAGAGAGAGTGTCCCATGTGACAGACT 1017
QY      341 ValGlyProTYrValProLeuTPRLeuGlyCyseGluAlaGlyPheTYrValaAspAlaGly 360
      1018 GTGGGCCCCCAAGTCCCTCTGCTGGCTGGGTGTGAGGACAGATTATATGCGATGGCGGA 1077
QY      361 ProProThrHleAlaPheThrProCyseGlyVhIleValCyseSerGluYseSerAlaIleTYr 380
      1078 CCCCCAATCAGCTCTTCAACCCCTGGGCGAGCTGTGTGCAAAAAGTCTGCCAAGTAC 1137
QY      381 TrpSerGlnIleProLeuProHleGlyThrHleAlaPheHleAlaIaCyseProPheCyse 400
      1138 TGGTCCAGATCCCACTGCCCCAGAGACCAAGCCGCTTATGCGGCTGTCCGTTCTGC 1197
QY      401 AlaThrGlnLeuValGlyGluGlnAenCyseIleYseLeuIlePheGlnGlyProIleAsp 420
      1198 GCCACGAGCTGTGTGTGAACGAAGTCAATGATTTTCCAGGTCCAGTGGAC 1257

```

```

RESULT 3
US-09-843-905A-3
; Sequence 3, Application US/09843905A
; Patent No. 6703487
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-843-905A-3

```

```

Alignment Scores:
Pred. No.: 9,89e-203 Length: 1257
Score: 1917.00 Matches: 342
Percent Similarity: 90.48% Conservative: 38

```

```

Best Local Similarity: 81.43% Mismatches: 38
Query Match: 83.71% Indels: 2
DB: 4 Gaps: 1
US-10-041-030-4 (1-420) x US-09-843-905A-3 (1-1257)
QY      1 MetPheSerProGlyGlnGluGlnHleScyValaProAenlyGluProValIySTYrGly 20
      1 ATGTTTTCCTCGATCAACAAAATCATAT-----CAATCAAGACCACTAAAATATGTGT 54
QY      21 GluLeuValIleuGlyTYrAenGlyAlaLeuProAenGlyAspArgGlyArgArgly 40
      55 GAACTCATTTCTTGAAGATTAATGATCTCTCCAAACGGTGATAGAGAAAGAGAGAA 114
QY      41 SerArgPheAlaLeuTYrIyAspArgProIyValaAenGlyValIyAspSerThrValHis 60
      115 AGTAGTTTGTCTTTGTTAAAGACTTAAGCAAAATGGGTGAAGCCGACCTGTCAT 174
QY      61 ValIleSerThrProGlnAlaSerIyValaIleSerCyseIyGlnHleSerIleSer 80
      175 ATGTCTTGATCTCTCAGGCTCAAGCAATAGCAACAAAGACAGATGACATATA 234
QY      81 TYrThrLeuSerArgAenGlnThrValValIleGluTYrThrHleAspIyAspThrAsp 100
      235 TATACCTTATCTCGGCCCCAGACTGTGTGTGATATATCTCATGACAGCAACCGAT 294
QY      101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIleThrAspThrIle 120
      295 ATGTTCAAGATTGGCCGGTGCAGTGAAGCCCATGATTTGTGATGTACTGCACAGGTT 354
QY      121 SerGlySerGlnAenThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
      355 CCTGGAAGTCMAAGTAATTCATGATACAGTCAAGTCAAGCACTATATCAAGATTGCC 414
QY      141 CyseArgIleValCyseAspArgAenGluProTYrThrAlaArgIlePheAlaIaGlyPhe 160
      415 TGCAGAAATCATATGTCAGAGCAATCTCTCTTACAGCAACGATTTATCTGAGGGTTT 474
QY      161 AspSerSerIyAspAenIlePheLeuGlyGluValaIleIleSTYrIyAspAenProAspGly 180
      475 GACTCATCAAAAACATCTTTCTTGGGAGAAAGGCTGCCAATAGAAACATCAGATGGA 534
QY      181 HlMeArbglYleuthrThraenglyValleuValMeChlProArnglyglYpHeThr 200
      535 CAGATGAGTGGCTTGAACCAATAGTGTCTTGTGATGATCATCCAGCAATGGGTTCA 594
QY      201 GluGlusErGlnProGlyValTPArGgJuileSeValCyseGlyAspValTYrThrLeu 220
      595 GAAAGCTCCAAAGCTGGAATATGAGAGAAATATCGTGTGTGGAAATGATATTACCTA 654
QY      221 ArgGluThrArgSerAlaGlnGlnArglyLyLeuValGluSerGluThraenValleu 240
      655 CGTGAACCCAGATCCGCTCAGCAGAGAGAGAAATGATGGAATGAACCAATCAGTTA 714
QY      241 GlnAerGlySerLeuIleAspLeuCyseGlyAlaThrLeuLeuTPArGThrAlaAspGly 260
      715 CAGAGTGGCTGTGTAATTAATCTGTGTGCAACATTTGTATGGGTGTGCAAGAGC 774
QY      261 LeuPheHleThrProThrGlnlyVshIleGluAlaLeuArnglInglIleAenAlaIa 280
      775 CTTTCCACACTCTCACTCGTGAAGCATTTAGAAAGCTTTAAGACAGAAATCAATGACGA 834
QY      281 ArgProGlnCyseProValGlyLeuAenThrLeuAlaPheProSerIleAenArglyGlu 300
      835 CGACCTCAAGTCCCTTGAAGTTCAACACATTCATTCATGTATGTAAGAGGAAAGC 894
QY      301 ValValGluGluIlyseGlnProTPArIleTYrLeuSerCyseGlyVhIleValHleGlyTYrHis 320
      895 GTTGTAGATGAAGAAACCAACATGCGTATATCTTAATGCGGCAATGATACATGCTATCAT 954
QY      321 AenTPRGlyVhIleArgSerAenThrGluAlaAenGluArnglIlyCyseProMetCyseArgThr 340
      955 AACTGGGAAGAAACAAAGAAAGTGAATGCGAAGATCGTATGTGTGATGCTCT 1014

```

```
QY      341 ValGlyProTyrValProLeuThrLeuGlyCYsGluAlaGlyPheTyrValAspAlaGly 360
      1015 GTTGGTCCCTAATGCTCTCTGTGGCTTGATGTAAGAGCTGGATTTTATGTGAAGCCGCC 1074
QY      361 ProProThrHisAlaPheThrProCYsGlyHisValCYsSerGluLysSerAlaLysTyr 380
      1075 CCTCCAAACCAATGCGTTTAAAGCCGCTGGGCAATGTGTGTTCAGAAAAAGCAACTGCCAT 1134
QY      381 TPSSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCYsProPheCYs 400
      1135 TGGTCCCAATGCCACTTCTCTCATGGTACTCATATTTTCATGACACCTGTCCCTTTGT 1194
QY      401 AlaThrGlnLeuValGlyGluGlnAsnCYsIleLysLeuIlePheGlnGlyProIleAsp 420
      1195 GCACATCAAGTTGGCTGTGTAACAAGGCTATCATCAAGCTTATTTTCAAGAGACTTAAAC 1254

RESULT 4
US-09-843-905A-1
; Sequence 1, Application US/09843905A
; Patent No. 6703487
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-843-905A-1

Alignment Scores:
Pred. No.:      3,55e-202      Length:      1257
Score:          1912.00      Matches:      341
Percent Similarity: 90.48%      Conservative: 39
Best Local Similarity: 81.19%      Mismatches: 38
Query Match:      83.49%      Indels:      2
DB:              4          Gaps:      1

US-10-041-030-4 (1-420) x US-09-843-905A-1 (1-1257)

QY      1 MetPheSerProGlyGlnGluHisCyAlaProAsnLysGluProValIlyrTyrGly 20
      1 AGGTTTCTCGATCAAGAAATATCAT-----CCTTCCAAAGCCCCAGTAAATATATGCG 54
DB
QY      21 GluLeuValValLeuGlyIlyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgGly 40
      55 GAACCTATTGTCTTAGATATATATGATCTCTCCCAACGGTGATGAGAGAGAGGAAA 114
DB
QY      41 SerArgPheAlaLeuTyrIlyrAspArgProLysAlaAsnGlyValLysProSerThrValHis 60
      115 AGTAGTTGCTTTGTTTAAAGACTTAAGCAATAGGGATGAGGCTTACACCGTGCAC 174
DB
QY      61 ValIleSerThrProGlnAlaSerIlyAlaIleSerCyAlaGlyGlnHisSerIleSer 80
      175 ATTGCAATGATCTCTCAAGCTGCCAAGCAATAGCAACAAGAGACCATATGATATCA 234
DB
QY      81 TyrThrIleSerArgAsnGlnThrValValValGlyTyrThrHisAspLysAspThrAsp 100
      235 TATACTTATCTCGAACCCAGACGGTGGTGAATATATCATGACAGCAACTGTAT 294
DB
QY      101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
      295 AGCTTTCAGATTGTGCGTCAACTGAAAGTCCATATTGTTTGTAGTAACTGACACCGTT 354
DB
QY      121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
```

```
DB      355 CCTGGAAGTCAGAGTAATTCGACACGCGACGATACAAAGACATATCAAGATTGCC 414
      141 CyAlaGlyIleValCYsAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaLysPhe 160
      415 TGTAGATCATATATGAGACGCGACGTCCCTTTACAGCTCGGATTTATGCTGACAGGTTT 474
QY      161 AspSerSerIlyAsnIlePheLeuGlyGluLysAlaAlaLysTyrLysAsnProAspGly 180
      475 GATTATCATAAAAACATCTTTCTTGGGAGAAAGCTGCCAAGTGAAGACATCTGATGGG 534
DB
QY      181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
      535 CAGATGATGAGCTTACACATATAGAGATTCTTGATGATGATCCACGTAATGGGTTTACA 594
DB
QY      201 GluGluSerGlnProGlyValIlyrArgGluIleSerValCYsGlyAspValTyrThrIleu 220
      595 GAAAGCTCCAAACCTCGAATATGAGAGAAATATCATGATGTGGGAAATGTCTTCACTCG 654
DB
QY      221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLysLeuValGluSerGluThrAsnValLeu 240
      655 CTGTAACCAATCAATCGCTCAGCAGAGAGAAAGATGGTGAATTTGAAACCAATACGTA 714
DB
QY      241 GlnAspGlySerLeuIleAspLeuCYsGlyAlaIleThrLeuLeuTyrArgThrAlaAspGly 260
      715 CAAGATGAGCTCCTTATATGACCTTTGTGTGCAACCTTGCTGGCTGCTGCTGCTGCTGCT 774
DB
QY      261 LeuPheHisThrProThrGlnLysHisIleLeuAlaLeuArgGlnGluIleAsnAlaAla 280
      775 CTTTCCCATATCTCTTACGTGTAAGCATTTAAAGCTTTTAAACAGAGATCAATGACACT 834
DB
QY      281 ArgProGlnCyProValGlyLeuAsnThrIleuAlaPheProSerIleAsnArgLysGlu 300
      835 CGCGCGAGTCCCTGTAGGCTTCAACATACCTTCCCAAGTATGAAGAGAGAT 894
DB
QY      301 ValValGluGluLysGlnProTyrPalatyrLeuSerCYsGlyHisValHisGlyTyrHis 320
      895 GTTGTAGATGAAGAACCAACATGGGTATATCTAACTCGGACCATGCTCATGTATATCAT 954
DB
QY      321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGluAlaGlyLysCyProMetCyAspArgThr 340
      955 AACTGGAGAAACAAAGAAAGAGCGTGAAGGCTGTGAATGTCTTATGTGTAAGTCT 1014
DB
QY      341 ValGlyProTyrValProLeuThrLeuGlyCYsGluAlaGlyPheTyrValAspAlaGly 360
      1015 GTTGGTCCCTAATGCTCTCTGTGGCTTGATGTAAGAGCTGGATTTTATGTGAAGCCGCC 1074
DB
QY      361 ProProThrHisAlaPheThrProCYsGlyHisValCYsSerGluLysSerAlaLysTyr 380
      1075 CCTCCAAACCAATGCGTTTAAAGCCGCTGGGCAATGTGTGTTCAGAAAAAGCAACGCGCTTAC 1134
DB
QY      381 TPSSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCYsProPheCYs 400
      1135 TGGTCCCAATGCCACTTCTCTCATGGTACTCATATTTTCATGACACCTGTCCCTTTGT 1194
DB
QY      401 AlaThrGlnLeuValGlyGluGlnAsnCYsIleLysLeuIlePheGlnGlyProIleAsp 420
      1195 GCACATCAAGTTGGCTGTGTAACAAGGCTATCATCAAGCTTATTTTCAAGAGACTTAAAC 1254

RESULT 5
US-09-843-905A-11
; Sequence 11, Application US/09843905A
; Patent No. 6703487
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; NUMBER OF SEQ ID NOS: 15
```

```

SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 1338
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (513)..(513)
OTHER INFORMATION: unsure
US-09-843-905A-11

Alignment Scores:
Pred. No.: 3,57e-172 Length: 1338
Score: 1642.00 Matches: 299
Percent Similarity: 81.38% Conservative: 42
Best Local Similarity: 71.36% Mismatches: 74
Query Match: 71.70% Indels: 4
DB: Gaps: 3

US-10-041-030-4 (1-420) x US-09-843-905A-11 (1-1338)
QY 3 SerProGlyGlnGluGlnHisCySAAProAmluPProVallyrGlyGluLeu 22
DB 88 TCTCCCGGT-----GAGATGGCGCAGCCAGGAGAGCCATCAAGTATGTTGAATC 141
QY 23 ValValleuGlyTyraAnglyAlaLeuProAnglyAAspArgGlyArgArgy 42
DB 142 ATGCTCTGGGCTTACATATGTTCTGGCAAGTGGGACAAGGCGCCGCGCAAGCGCG 201
QY 43 PheAlaLeuTyraPAspArgProlySAAlaAnglyVallyrProSerThrValHisValIle 62
DB 202 CTGGCACTGAGCGCGCGGTGCGACGCCAGCGGAGTGAAGCAAGACGTATGACACCATC 261
QY 63 SerThrProGlnAlaSerLySAAlaIleSerCySAleGlyGlnHisSerIleSerTyThr 82
DB 262 TCCACGCGCGCTCGCTCTCAAGCAGCATGATACCGTGTGACACACCATCTGTATACA 321
QY 83 LeuSerArgAnglnThrValValGlyTyThrHisAspLyAspThrAspMetPhe 102
DB 322 CTGTCCCGGAGCAGCATCGGTCAATAGTGAATATACATATGATGCGCACAGACATGTC 381
QY 103 GlnValGlyArgSerThrGlySerProIleAspPheVallyrThrAspThrIleSerGly 122
DB 382 CAGATTGGCGCGCTCCACAGAAACATGATTGACTCTGTGTAAACAGACAGTCCCTGGA 441
QY 123 SerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAlaCySAArg 142
DB 442 GGA---GGGCGTCCGAGGCGCCCTTCTGCCACAGACCATCTCCGCTATGCTGCCGC 498
QY 143 IleValCySAAspArgAnglnProTyThrAlaArgIlePheAlaAlaGlyPheAspSer 162
DB 499 ATCTCTGTGACCGNCGCGCACCATTAATGCGCGCATCTATGCGCTGCGTCTGATGCC 558
QY 163 SerLyAsnIlePheLeuGlyGlyValAlaAlaLyThrLyAsnProAspGlyHisMet 182
DB 559 TCTAGCAACATCTCTCTGAGAGCGAGCGCCAAATGGCGGACCCCAAGATGCGCTGATG 618
QY 183 AspGlyLeuThrThrAnglyValLeuValMetHisPAspArgGlyGlyPheThrGlnGlu 202
DB 619 GATGAGACTGACCAATGAGTCTCTGTATGATGACCCGCGCAGCGGCTCTCCGAGGAC 678
QY 203 SerGlnProGlyVallyrPArgGlyLysSerValCySAIlyAspVallyrThrLeuArgGlu 222
DB 679 TCAAGCCCGCGGTCTGGCGGAGATCTCGGTCTGTGGAGATGTACATATGCGGAGAC 728
QY 223 ThrArgSerAlaGlnGlnArgGlyLyLeuValGlySerGlnThrAsnValLeuGlnAsp 242
DB 739 AGCGGCTTCAAGCCAGCGGCGGAGCATGTTGAAACAGTCAACGCTGCGACGAGAC 798
QY 243 GlySerLeuIleAspLeuCySAIlyAlaThrLeuLeuThrAlaAspGlyLeuPhe 262
DB 799 GGCTCTCTATCGACCTGTGTGGGCGCACACTGTGTGGGCGACACCGGCGGCGCTGCTG 858

```

```

QY 263 HisThrProThrGlnLyHisIleGlnAlaLeuArgGlnGluIleAsnAlaAlaArgPro 282
DB 859 CGGGCTCCACACTGAAAGCACTGAGGCGCCAGCGGACAGACCAATACAGCGCGCCC 918
QY 283 GlnCySAProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLySAValVal 302
DB 919 CAGTCCCGCTGGGCTTCAAGCATCTGAGCTTCCCAAGCCAGCCGCTGGCGCACAGG 978
QY 303 GlnGluLyS---GlnProTrrPalArgLyLeuSerCySAIlyHisValHisGlyTyHisAsn 321
DB 979 CCCGACAAACAGACAGCGCTGGGTCTACGTCGCTGGGAGGACATGCTCATGATGACAGCG 1038
QY 322 TrpGlyHisArgSerAspThrGlnAlaAnglyGluCySAProMetCySAArgThrVal 341
DB 1039 TGGGCTGCGCGGAGGAGGCGGCGCCAGAGAGCGGAATGTCCTCTGCGCGCTTGTG 1098
QY 342 GlyProTyValProleuThrPleuGlyCySAAlaGlyPheTyrValAspAlaGlyPro 361
DB 1099 GGGCTTATGTGCTCTTATGCTTGGCCAGAGGCGCGCTGCTGCTGACCCCTGGGCGG 1158
QY 362 ProThrHisAlaPheThrProCySAIlyHisValCySAArgLyAspSerAlaLyTyTrp 381
DB 1159 CTAAGCATGCTTTCACCTTGGCGGACAGTCTGCTGTGAAGAAGCTCCGCTACTGG 1218
QY 382 SerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCySAProMetCySAAla 401
DB 1219 GCCACACACACTGCGCCAGCGGACCCATGCTTTCAGAGCGCGCTGCTTGGCGG 1278
QY 402 ThrGlnLeuValGlyGlnGlnAsnCySAIlyLeuLeuPheGlnGlyProIleAsp 420
DB 1279 GCCTGCTTACCGGAGCATGAGTCTGCTGCGCTCATTTTCAAGGCGCGCTGAT 1335

RESULT 6
US-08-136-277-1
Sequence 1, Application US/08136277
Patent No. 5644045
GENERAL INFORMATION:
APPLICANT: MANDEL, Jean-Louis
APPLICANT: AUBOURG, Patrick
APPLICANT: MOSSER, Jean
APPLICANT: SARDE, Claude
TITLE OF INVENTION: X-LINKED ADRENOLEUKODYSTROPHY GENE AND
TITLE OF INVENTION: CORRESPONDING PROTEIN
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSER: Young & Thompson
STREET: 745 South 23rd Street
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/136,277
FILING DATE: 15-OCT-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: B2272
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2750 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

```

```

/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 387..2624
/ US-08-136-277-1

Alignment Scores:
Pred. No.: 0.00763 length: 2750
Score: 120.50 Matches: 104
Percent Similarity: 33.94% Conservative: 63
Best Local Similarity: 21.14% Mismatches: 155
Query Match: 5.26% Indels: 171
DB: 1 Gaps: 26

US-10-041-030-4 (1-420) x US-08-136-277-1 (1-2750)

QY 2 PheSerProGlyngIngluHisCySaLaProAsnLysGluProValIysTyrgIu 21
DB 1167 TTTCTCGCCC-----AAGTTCGGGGAG 1187

QY 22 LeuValValLeuGlyTyraEnglyAlaLeuProAsnGlyAspArgGlyYrArgLysSer 41
DB 1188 CTGGTG-----GCAGAGAGGGCGGGCGGAAGGGG 1217

QY 42 ArgPheAla---LeuTyriYsarqProLySaLaAsnGlyValIysProSerThrValHis 60
DB 1218 GAGCTGGCTACATGACCTCGCTGCTGCGCCAC----- 1253

QY 61 ValIleSerThrProGlnAlaSerLySaLaIleSerCySaLyGlyGlnHisSerIleSer 80
DB 1254 -----TCGAGAGAGATCGCTTCTATGCGGGCCATGAGTGGAG 1292

QY 81 TyrThrLeu-----SerArgAsnGlnThrValVal 91
DB 1293 CTGGCCCTCTACAGGCTCTCTCAACAGACCTGGCTCGCAGATCAACCTCATCTTCG 1352

QY 92 Glu-----TyrThrHisAspLySaAspThrAspMetPheGlnValGlyArqSerThr 108
DB 1353 GAAAGCTGTGATGATTTATGCTGAGACAGATTCTCATGAGATGTGCGAGCGCTCG 1412

QY 109 GluSerProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThrAspGlu 128
DB 1413 GGCCTGCTCATGTGCTGCTGCCCATCATCATCCACCTGCTACTCATGACATGACA 1472

QY 129 AlaGlnIleThrGlnSerThrIleSerArgPheAlaCySaArgLysValCySaAspArgAsn 148
DB 1473 GAGGCGGTGAAGAGGAGCGCTTGGAAAAGAGAGAGAGAGAGCTGTGAGCGCAGACA 1532

QY 149 GluProTyThr-----AlaArgIlePheAlaAlaGlyPheAspSerSerLySaAsnIle 166
DB 1533 GAAAGCTTCACTATTCGCCGCAACCTTCGACAGCGGGCTGCAATGCCATTGACGGATC 1592

QY 167 -----PheLeuGlyGluValAlaAlaLyTrpLysAsnProAspGlyHisMetAsp 183
DB 1593 ATGCTCTCGTACAA--GGAGGTGACGAGCTGCTGCTACACAGCCCGGGTGCACGAGAT 1651

QY 184 -----GlyLeu----- 185
DB 1652 GTTCCAGGATTTTGAAGATGTTACGGCTGTCACTTCAAGAGGCCAGGAGCTAGAGA 1711

QY 186 ThrThrAsnGlyValLeu-----ValMetHisProArgGlyGlyPheThrGlu 201
DB 1712 CGCTCAGGGGGGTCTGGAGCATAGAGCGGTGTGTCCTCGTGGAGGGCCCTGAA 1771

QY 202 GluSerGlnProGlyValIysTrpArgGluIleSerValCyGlyAspValIysThrLeuArg 221
DB 1772 ATCCAGGCGCAGGT-----GGTGAATGGGAAACAGGGGATCATCTCGCA 1816

QY 222 Glu-ThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnVal---Ile 240
DB 1817 GAAATATCCCATCGTACAGCGCTCAGAGAGAGGTGTGTGGCCAGGCTTCAACATCAGGGT 1876

```

```

QY 240 uGlnAspGly-----SerLeuI 246
DB 1877 GGAGGAAGGACATGCATCTGCATACAGAGCCCAATGGCTCGCGCAGAGCTCCCTGT 1936

QY 246 eAspLeuCySglYalathrLeuLeuTrpArgThrAlaAspGlyLeuPheHisThrProTh 266
DB 1937 CCGATCTCGGGGTGG-----CTCTGGCCCACTGACGATGGTGTGTCTCTACMACCCCC 1990

QY 266 rGlnIysHisIleGluAlaLeuArgGlnGluIleAsnAlaAlaAspProGlnCyProVa 286
DB 1991 ACCCAAGGACATGTTCTTACATCCGAG-----AGGCCCTACATCTCTGT 2035

QY 286 lGly-----LeuAsnThrLeuAlaPheProSerIleAsnArgLySgluValGluG 304
DB 2036 GGGCTCCCTCGGTGACCAAGTATATACCCG-----GACTAGTGGAGGA 2080

QY 304 uLySglnProTrpAlaTy-----LeuSerCySglYHisVa 316
DB 2081 CATGCMAAGAAAGGCTACTCGAGACAGACCTGGAAGCCATCTCGAAGTGTGCACCT 2140

QY 316 lHis-----GlyTyHis-----AsnTrpGlyHis---Ar 325
DB 2141 GCACCAATCTCTGCAGCGGAGAGAGGTTGGAGGCTATGTGTACTGAAAGACSTCT 2200

QY 325 sSerAspThrGlu-AlaAsnGluArgGlyCySProMetCySaArgThrValGlyProTyrv 345
DB 2201 GTCCGGGTGCGAAGAGAGAAATCGGATGGCCGATGTTCTTACCAAGGCCCAAGTA 2260

QY 345 aProLeuTrpLeuGlyCySgluAlaGlyPheTyValAspAlaGlyPro----- 361
DB 2261 CCCTCTCTGTGAATGACACACGCG-----CCGTGAGCATGGA 2299

QY 362 -----ProThr 364
DB 2300 CGTGAAGCAAGATCTTCCAGCGGCCAAGAGCGGGCATTCCTGCTCTCATCAC 2359

QY 364 lAslaPheThrProCySglY-----HisValCyS-----SerG 375
DB 2360 CCACGGGCGCTCCCTGTGGAAATACACACACTTGTCTACATGATGGAGAGGGCGG 2419

QY 375 lulySerAlaLyTyTrpSerGlnIlePro 385
DB 2420 CTGGAAGTTCGAGAGCTGACTGACGCGCC 2451

RESULT 7
US-08-479-403-1
/ Sequence 1, Application US/08479403
/ Patent No. 5869039
/ GENERAL INFORMATION:
/ APPLICANT: MANDEL, Jean-Louis
/ APPLICANT: AUBOURG, Patrick
/ APPLICANT: MOSSER, Jean
/ APPLICANT: SARDE, Claude
/ TITLE OF INVENTION: X-LINKED ADRENOLEUKODYSTROPHY GENE AND
/ NUMBER OF SEQUENCES: 23
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Young & Thompson
/ STREET: 745 South 23rd Street
/ CITY: Arlington
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22202
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/479,403
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:

```

```

1 NAME: PATCH, Andrew J.
2 REGISTRATION NUMBER: 32,925
3 REFERENCE/DOCKEN NUMBER: B22722D1V
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: 703-521-2297
6 TELEFAX: 703-665-0573
7
8 TELEX: 248425 EMBON
9
10 INFORMATION FOR SEQ ID NO: 1:
11 SEQUENCE CHARACTERISTICS:
12 LENGTH: 2750 base pairs
13 TYPE: nucleic acid
14 STRANDEDNESS: single
15 TOPOLOGY: linear
16 MOLECULE TYPE: cDNA
17
18 FEATURE:
19 NAME/KEY: CDS
20 LOCATION: 387..2624
21
22 US-08-479-403-1

```

| | | |
|------------------------|-------------------|------------------|
| | Alignment Scores: | |
| Pred. No.: | 0.00763 | Length: 2750 |
| Score: | 120.50 | Matches: 104 |
| Percent Similarity: | 33.94% | Conservative: 63 |
| Best Local Similarity: | 21.14% | Mismatches: 155 |
| Query Match: | 5.26% | Indels: 171 |
| DB: | 2 | Gaps: 26 |

US-10-041-030-4 (1-420) X US-08-479-403-1 (1-2750)

| | | | |
|----|------|---|------|
| Oy | 2 | theserProGlyGlnGluGlnHieCysValaProAnlyvGluProValIlePtyGlyGlu | 21 |
| Db | 1167 | TTCTCGCC-----AAATTGGGGAG | 118 |
| Oy | 22 | LeuValValLeuGlyTyraSnglyAlaLeuProAnGlyAsaArgGlyAraArgPysSer | 41 |
| Db | 1188 | CTGGTG-----GCAGAGAGAGCGCGGGAAGGGG | 121 |
| Oy | 42 | ArgPheAla---LeuTyTyLeaArgProGlyAlaAnGlyValIlePseroSerThraHis | 60 |
| Db | 1218 | GAGCTGCCTCATGCACTCGGTGTGGTGGCAAC----- | 125 |
| Oy | 61 | ValIleSerThrProGlnAlaSerIlyAlaIleSerCylvSgLYGlnHieSerIleSer | 80 |
| Db | 1254 | -----TCGAGAGAGATCGCCTTCTATGGGGGCCATAGAGTGAAG | 129 |
| Oy | 81 | TyrThrIleu-----SerArgAnGlnThrValValVal | 91 |
| Db | 1293 | CTGGCCCTCGTACAGCGCTCTACAGAGACGCGCTGCAGATCAACTCATCTCTCTCG | 135 |
| Oy | 92 | Glu-----TyrThrHieAsrlyuAerThrIleAsrMetPheGlnValGlyArgSerThr | 108 |
| Db | 1353 | GAAAGCCTGGGTATGTTATGCTGTGAAGAGTTCCTCATGAAGTATGTGTGGAGCCCTCG | 141 |
| Oy | 109 | GluSerProIleAsrPheValValThrAspThrIleSerGlySerGlnAsnThrAspGlu | 128 |
| Db | 1413 | GAGCCTGCATAGTGGCTGTGCCATCATCACTGCCACTCGTACTCAAGAGTCAATGCA | 147 |
| Oy | 129 | AlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAspArgAsn | 148 |
| Db | 1473 | GAGCGCTGAAGAAAGCAGCCTTGGAAGAAAGAGAGAGAGACTGTGGCGAGCAGCA | 153 |
| Oy | 149 | GluProTyThr-----AlaArgIlePheAlaIaGlyPheAsrSerSerIlyAsnIle | 166 |
| Db | 1533 | GAAAGCTTCATATTGCCCCGCAACTCTCTGACAGAGCGCTGCAGATGCCATTGAGCGGATC | 159 |
| Oy | 167 | -----PheLeuGlyGlnlyuAlaIaIyATPlyvAsnProAspGlyHieMetAsr | 183 |
| Db | 1593 | ATGTGCTGTCACAA--GGAAGTGAAGAGCTGCTGAGTACACAGCCGGGTGACAGAGAT | 165 |
| Oy | 184 | -----GlyLeu----- | 185 |
| Db | 1652 | GTTCCAGGATATTGAAGATGTTCAAGCGCTGCTCATTTCAAGAGCGCCAGGAGCTAGAGAG | 1711 |

| | | | |
|----|------|---|------|
| Oy | 186 | ThrThrnnglValIleu-----ValmethisProAsglyPheThrglu | 201 |
| Db | 1712 | CGCTCAGCGCGGGCTCTGGACACATAGGCCGCTCTGGTCTCCGTGGAGGGCCCCCTGA | 1771 |
| Oy | 202 | GluserglnProgluValTTPArgluIleSerValCysgluAerValTyThrleuArg | 221 |
| Db | 1772 | GATCCGAGGCCAGGT-----GATGAGTGTGGAGAGGGGATCATCTTGGCA | 1818 |
| Oy | 222 | Glu-ThrArgserAlaInglnArglyLysIleValGluserglnThrAerVal--Le | 240 |
| Db | 1817 | GAAACATCCCATCGTCAAGCCCTCAGAGAGAGTGTGTGGCCAGCTCAACATCAGGT | 1878 |
| Oy | 240 | uglnAerGly-----SerIleuI | 246 |
| Db | 1877 | GGAGAGAGGCATGCATCTCATCACAGGCCCAATGCTGGGGCAAGACTCCCTGT | 1936 |
| Oy | 246 | eAerPleuCysgluAlaThrleuLeuTTPArgThrAlaAerGlyLeuPheNisThrProth | 266 |
| Db | 1937 | CCGAGTCTCGGTGGTGG-----CTCTGGCCCACTCAGTGTGTGCTTCAAGCCCC | 1990 |
| Oy | 266 | rglnLysenIstleGluAlaAerArgInglnIleAsnAlaAlaThrProInglyProVa | 286 |
| Db | 1991 | ACCCAGGACATGTTCTTACATCCCGCA-----AGCCCTTCAATGCTGT | 2035 |
| Oy | 286 | IglY-----LeuAsnThrleuAlaPheProSerIleAsnArgLysgluAlaInglnI | 304 |
| Db | 2036 | GGGCTCCCTGGTACCAAGCATGATCTACCG-----GACTCAGTGGAGGA | 2080 |
| Oy | 304 | uLysglnProTTPAlaTy-----LeuSerCysgluNisIleVa | 316 |
| Db | 2081 | CATCAAAAGAGAGGCTACTCGAGACAGACCTGAAGCCATCTGGACGTGTGCACCT | 2140 |
| Oy | 316 | IHis-----GlyTyNis-----AerTTPArgIHis--Ar | 325 |
| Db | 2141 | GCACACATCTCGACGCGGAGGAGAGGTTGGAGAGCTATGTGTACTGGAAAGACTCT | 2200 |
| Oy | 325 | gSerAerPThrglu-AlaAsnGluArgLysCysPrometCysArgThrValIslProTyv | 345 |
| Db | 2201 | GTCGGGTGGCAGAGACAGAGATCGGATGGCCCGCATGTTTACACACAGGCCCAAGTA | 2260 |
| Oy | 345 | alProleuTTPleuGlyCysgluAlaGlyPheTyValAerAlaGlyPro----- | 361 |
| Db | 2261 | CGCCCTCTGTGATGAATGACACACAG-----CCGAGACATGCA | 2299 |
| Oy | 362 | -----ProThn | 364 |
| Db | 2300 | CGTGAAGGCAAGATCTTCCAGGGGGCCAGAGACGGGGCATTTGCCCTGTCTCATAC | 2355 |
| Oy | 364 | IslaIaPheThrProCysglY-----HisValCys-----SerG | 375 |
| Db | 2360 | CCACCGGCGCTCCCTGTGGAAATGCACACACACTTCTACATTCATGGAGGAGCGG | 2419 |
| Oy | 375 | IuLysSerAlaLysTyTTPSerIlnIlePro | 385 |
| Db | 2420 | CTGGAAGTTCAGAGACTGCACTTACGCTGCC | 2451 |

```

? RESULT 8
? US-08-835-734-1
? Sequence 1, Application US/08835734
? Patent No. 6013769
? GENERAL INFORMATION:
? APPLICANT: MANDEL, Jean-Louis
? APPLICANT: AUBOURG, Patrick
? APPLICANT: MOSSER, Jean
? APPLICANT: SARDE, Claude
? TITLE OF INVENTION: X-LINKED ADRENOLEUKODYSTROPHY GENE AND
? TITLE OF INVENTION: CORRESPONDING PROTEIN
? NUMBER OF SEQUENCES: 23
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Young & Thompson
? STREET: 745 South 23rd Street
? CITY: Arlington
? STATE: VA
?

```

```

COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,734
FILING DATE:
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/479,403
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: B2722DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2750 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 387..2624
US-08-835-734-1

```

```

Alignment Scores:
Pred. No.: 0.00763 Length: 2750
Score: 120.50 Matches: 104
Percent Similarity: 33.94% Conservative: 63
Best Local Similarity: 21.14% Mismatches: 155
Query Match: 5.26% Indels: 171
DB: 3 Gaps: 26

```

US-10-041-030-4 (1-420) x US-08-835-734-1 (1-2750)

```

QY 2 PhasEPrGlyInGluGluHisCysAlaProAsnGlyLeuProValIysTyrGlyGlu 21
DB 1167 TTCTCGCCC-----AAGTTCGGGGAG 1187
QY 22 LeuValIalLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgAspSer 41
DB 1188 CTGGT-----GCAGAGAGGCGCGCGCGAGGGG 1217
QY 42 ArgPheAla--LeuTyrIysArgProLysAlaAsnGlyValIysProSerThrValHis 60
DB 1218 GAGCTCGCTACATGACATCGCTGTCGTGGCCAAAC----- 1253
QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysAlaGlyGlnHisSerIleSer 80
DB 1254 -----TCGAGAGAGATCGCTTCATGTCGGGCGCATGAGTCGAG 1292
QY 81 TyrThrLeu-----SerArgAsnGlnThrValVal 91
DB 1293 CTGGCCCTGCTACAGGCTCTCAACGAGACTGGCTTCGACATCAACTCATCTTCG 1352
QY 92 Glu-----TyrThrHisAspLysAspThrAspMetPheGlnValGlyArgSerThr 108
DB 1353 GAACGCTGTGATGTATGTCGAGCACTTCCTCATGATGATGTGAGAGCGCTCG 1412
QY 109 GluSerProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThrAspGlu 128
DB 1413 GGCCTGCTCATGTGCTGCCCATCATCATCGCCACATGCTACTCAAGAGTCAGATGCA 1472
QY 129 AlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAspArgAsn 148

```

```

DB 1473 GAGGCGTGAAGAAGGCAAGCTTGAAAAAGAGAGAGAGCTGTGAGGAGGCA 1532
QY 149 GluProTyrThr-----AlaArgIlePheAlaAlaGlyPheAspSerIysAsnIle 166
DB 1533 GAAGCTTCACTATTATGCCCGCAACTCTGACAGCGGTGTCAGATGCCATTGACGGATC 1592
QY 167 -----PheLeuGlyGluLysAlaIleLysTyrLysAsnProAspGlyHisMetAsp 183
DB 1593 ATGTGCTGTACAA--GGAGGTGACGAGCTGGCTGCTACACAGCCCGGTGCACAGANT 1651
QY 184 -----GlyLeu----- 185
DB 1652 GTTCAAGGATATTGAAGATGTTCAAGCGTTCATTCAAGAGCCCGAGGACTAGAGA 1711
QY 186 ThrThrAsnGlyValLeu-----ValMetHisProAspGlyGlyPheThrGlu 201
DB 1712 CGCTCAGCGGGGTCTGGAGACCATGAGCCGGTCTGTGTCGTGTCGAGAGGCCCTTAA 1771
QY 202 GluSerGlnProGlyValThrArgGluIleSerValCysGlyAspValTyrThrLeuArg 221
DB 1772 GATCCGAGGCGAGT-----GGTGAATGTGAACAGGGGATCATCTGCGA 1816
QY 222 Glu-ThrArgSerAlaGlnAlaArgGlyLysLeuValGluSerGlnThrAsnVal---Le 240
DB 1817 GAACATCCCATTCCTCAGAGCCCTCAGAGAGGTGTGTGGCCACCTCAACATCAGGGT 1876
QY 240 uGlnAspGly-----SerLeuI 246
DB 1877 GGAGGAGGCGATGATCTGCTCATCATCAGAGCCCATGAGCTGGCGCAAGCTCCCTGT 1936
QY 246 eAspLeuCysGlyAlaThrLeuLeuThrArgThrAlaAspGlyLeuPheHisThrProTh 266
DB 1937 CCGAATCTGGGTGGG-----CTCTGGCCACGAGAGGTGTGTGCTCTCAAGCCGCC 1990
QY 266 rGlnIleHisIleGlnAlaLeuArgGlnGluIleAsnAlaAlaArgProGlnCysProVa 286
DB 1991 ACCCGAGGCGCATGTTCTACATCCGCGAG-----AGGCCCTACATGTCGT 2035
QY 286 lGly-----LeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValValGluI 304
DB 2036 GGGCTCCCTGCGTGCAGAGATGATCTACCCG-----GACTCAGTGAAGA 2080
QY 304 uLysGlnProThrAlaTyr-----LeuSerCysGlyHisVa 316
DB 2081 CATGCAAAAGAAAGGCTACTCTGAGAGAGACCTGGAAGCATCTGAGCTGTGCACCT 2140
QY 316 His-----GlyTyrHis-----AsnTyrGlyHis---Ar 325
DB 2141 GCACCAATCTCGACGGGAGGAGGTGGAGGCTATGTGTGACTGGAAGAGAGTCTCT 2200
QY 325 gSerAspThrGlu-AlaAsnGluArgGluCysProMetCysArgThrValGlyProTyrV 345
DB 2201 GTGGGTGGCGGAGAGAGAGATGCGATGCGCCGATGTTCTTCCACAGGCCCAAGTA 2260
QY 345 aLProLeuThrLeuGlyCysGluAlaGlyPheTyrValAspAlaGlyPro----- 361
DB 2261 GCGCTCTCGATGAATGACACAGCG-----CCGTGAGCATGCA 2299
QY 362 -----ProThrH 364
DB 2300 CGTGAAGGCAAGATCTTCCAGGGCGCAAGAGCGGGGCAATGCGCTGTCTCATCAC 2359
QY 364 lAlaPheThrProCysGly-----HisValCys-----SerG 375
DB 2360 CCACCGGCGCTCTCGTGAATAATACACACACACTTCTACAGTTGATGGGAGGGCGG 2419
QY 375 LuLysSerAlaLysTyrThrSerGlnIlePro 385
DB 2420 CTGGAAGTTGAGAAAGCTGACTCGTGGCC 2451

```

RESULT 9
US-09-103-840A-2/c

[illegible]

Pred. No.: 7.43e+03 Length: 4411529
 Score: 113.50 Matches: 98
 Percent Similarity: 34.028 Conservative: 51
 Best Local Similarity: 22.37% Mismatches: 161
 Query Match: 4.96% Indels: 129
 DB: 3 Gaps: 21

US-10-041-030-4 (1-420) x US-09-103-040A-1 (1-4411529)

```

QY      4 ProgluInglu-----GluHisCybAlaProAmluProValuLeuTyr 19
DB      1251307 CCTGGGCAACAAGGCGACGATCGATGCGCCGCGACCATCTGCC----- 1251257

QY      20 GlyValuLeuValValleuGlyTyrAenGlyAlaLeuProAenGlyAspArgGlyArg 39
DB      1251256 GGTGCCCGCTGCATTTGATTACAC-----CGAGGTGGCGCT 1251218

QY      40 LysSerArgPheAlaLeuTyrLysArgProLysAlaAenGlyValLysPro----- 56
DB      1251217 -----GCGCTATCTGCGCGCGCGGATCCAGATCGCGCGATCGGGATGA 1251170

QY      57 -----SerThrValHisValLleSerThrPro 65
DB      1251169 CTTGCTGATCGGATGATGCTGCTGGAAACAGAAACCTGTGCAGTTCGCGCTGGCCC 1251110

QY      66 GluAlaSerLysAlaLleSerCybLysGlyGlnHisSerLle-----Ser 80
DB      1251109 GAATCGCGCTGCTGTGACGCTGCGCTCGGGAGCCACCAGTCGATGCCGCGATATGGA 1251050

QY      81 TyrThrLysSerArg-AsnGlnThrValValGlu----- 92
DB      1251049 CACGATTCCTGACGGGGGATACCGACGCGGTGCAACAGATCTTCCAGCGCGGTGC 1250990

QY      93 -----TyrThrHisAsp-LysAspThrAspMet---PheGlnValGlyArgSerThr 109
DB      1250989 CCGCCACTTCGCCACGACACGACGCGGATTCACAGCATGTATCGCGCATGTCGG 1250930

QY      109 LysEProLleAspPheValValThrAspThrLleSerGlySerGlnAenThrAspGlu- 128
DB      1250929 CAGCGCGGTCCCATTTACCGACGACCCGTATGCCCGCGACGACGAGCGGT 1250870

QY      129 -----AlaGlnLleThrGlnSerThr- 135
DB      1250869 ATTGCGCTCTGGAATCACTCGCGCGCTTGCCTATGCCGACATCAAGCGCTGACGA 1250810

QY      136 -----LleSerArgPheAlaCybArgLleValCybAspArgAsnGluProTyrThrA 153
DB      1250809 CCAGGCGGCTACGCGGTGTGTGTGACGATGATTCGGCAACTCAGCG----- 1250757

QY      153 LArgLlePheAlaLaglyPheAspSerSerLysAsnLlePheLeuGly-----GluL 171
DB      1250756 --GAACTCG-GAGCGCGGGGTGTCTCTCAAGCAACCGGATCAACCGGAGCCAAAGAC 1250700

QY      171 ySAlaLalAyETPrLysAsnProAspGlyHisMetAspGlyLeuThrThrAenGlyVal 191
DB      1250699 CGTTGGCGGAATGCGTA---CCCGGCGCCAAAGGCAACAGCACTG---GTGTCCGATTGG 1250646

QY      191 euValMetHisProArgGlyGlyPheThrGlnGlnSerGlnProGlyValTyrArgGly 211
DB      1250645 TCGCGGTGCGACCGATCACCGGCTGCGGACAG---CAGCGCGCGCTCTAGCAAC 1250589

QY      211 leSerValCybGlyAspValTyrThrLeuArgGluThrArgSerAlaGlnGlnArgGly 231
DB      1250588 TG-----ATGCGCGCGCGCGCGCATGTTGGAGCCCATGGCGAG 1250553

QY      231 yLeuValGluSerGluThrAenValLleGlnAspGlySer-----LeuLleAspLeuc 249
DB      1250552 GCGAGATCCAGAGATCGAATTCACCGTGGAGAGACAGCCAGCTGTGCTGACAGACC 1250493

QY      249 ySGLyAlaThrLeuLeuTyrArgThrAlaAspGlyLeuPheHisThrProThrGlnLys 269
DB      1250492 GCGGGGCGGAA-----CGCTGCGCGCAAGCGCGCGCTGCGCAACTGCGCAACTGC 1250442
  
```

```

QY      269 iSllleGluAlaLeuArgGlnGluLleAsnAlaAlaArgProGlnCybProValGlyLeuA 289
DB      1250441 ATCAGAGAGACTCATTCAGACGACCCGACACTGCGCGGTACTCTCGACGCATTCG 1250382

QY      289 enThrLeuAlaPheProSerLle-AsnArgLysGluValGlnGluLysGlnProTyr 308
DB      1250381 AGACTTCTGTCAGCGCGTGTGCGGACGAGAA----- 1250349

QY      309 AlaTyrLysSerCybGlyHisValHisGlyTyrHisAsnTPrGlyHisArgSerAspThr 328
DB      1250348 -----CAGCGTGGCTGCACCGC----- 1250331

QY      329 GluAlaAenGluArgGluCybProMetCybArgThrValGlyProTyrValProLeuTyr 348
DB      1250330 ---TTTGGCCAAAGGCTGCGCGCTTCCCGCGCGGTGTATCC----- 1250289

QY      349 LeuGlyCybGluAlaGlyPheTyrValAspAlaGlyProProThrHisAlaPheThrPro 368
DB      1250288 -----GGACCGCTTACACCGAGGTGACGAG 1250262

QY      369 CybGlyHisValCybSerGluLysSerAlaLysTyrTTPSerGlnLle 384
DB      1250261 CGTTGAGCTGCGGACCGGCGCGCGGTATCTTGTGCGATC 1250214

RESULT 11
US-09-252-991A-14397/C
/ Sequence 14397, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT FILING DATE: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 14397
/ LENGTH: 1590
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14397

Alignment Scores:
Pred. No.: 0.0591 Length: 1590
Score: 109.00 Matches: 107
Percent Similarity: 30.08% Conservative: 47
Best Local Similarity: 20.90% Mismatches: 177
Query Match: 4.76% Indels: 182
DB: 4 Gaps: 25

US-10-041-030-4 (1-420) x US-09-252-991A-14397 (1-1590)

QY      24 ValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArg-----GlyArgArgLys 40
DB      1477 CTGCTCGCGCGGCTCATATCGGTTCCTCA---GGCGAAAGCCAGCTCGCGGTGCGGAACG 1421

QY      41 SerArgPheAlaLeuTyrLysArg-----ProLysAlaAenGlyValLysProSer 57
DB      1420 CTTGCTGCGGAGCTCATATCATCGGTTCGGAATTCACCTGCGCACATCGATCGCTGACCGGA 1361

QY      58 ThrValHisValLleSerThrProGlnAlaSerLysAlaLleSerCybLysGlyGlnHis 77
DB      1360 ACTGCGGCGCGCTTCGAGATCTTCGCTGCTGCGGCGGCGCTGATCAGCCGAT 1301

QY      78 -----SerLysTyrThrLysSerArg----- 85
DB      1300 CCACCATGACTGCTGCGCGCGAGATCCTTGGCGGAGGTATCGCTGTCCGACAC 1241

QY      86 -----Asn 86
  
```


| | | | | | |
|----|------|--|--|-----------------------------------|------|
| QY | 142 | ----- | ArglleVal | CysAspArgAsnGluProTyrThrAlaArgIle | 155 |
| | | | | | |
| Db | 925 | CAGGCAATTGGCCCTGGGGCGGGAGTCTCTCAAAAGAAAGAGCCCTGGGGCTGTGGGGCG | | | 984 |
| QY | 156 | Phe----- | AlaIaGlyPheAspSerSerLysAsnIlePheLeuGly | 169 | |
| | | | | | |
| Db | 985 | CTGCTTCAGAGAGAGACCCCGGACCCCTTCTGAGCAGTGCAGAGAACATTGCTGTGGGG | | | 1044 |
| QY | 170 | ----- | GluLysAlaAlaLysTrpLysAsn----- | ProAspGlyHisMetAsp | 183 |
| | | | | | |
| Db | 1045 | CTTGCAACAGAGAAAGCCTGTGCTTGGCTGTGACGCCAACATCAGCAAGTGGAT--- | GAG | 1101 | |
| | | | | | |
| QY | 184 | GlyLeuThrThrAsnGlyValIleuValIleMetHisProArgGlyGlyPheThrGluGluSer | 203 | | |
| | | | | | |
| Db | 1102 | AGTGGCGGTGTGAATGGCTGTGATATACCAATGATGCTTACGAGAGC----- | AGT | 1148 | |
| QY | 204 | GlnProGlyValAlaTrpArgGluIleSerValCysGly----- | AspValTyrThrLeu | 220 | |
| | | | | | |
| Db | 1150 | CGCCCACTATTTGGAGACAAATTTACAAATATCTTATGTGGAGAGACGTGCTCTCTTG | | 1208 | |
| QY | 221 | ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu | 240 | | |
| | | | | | |
| Db | 1210 | GCCCGTGGGGCCACCGGAGCCCTGACGAGGAGAGTCTCCCAAGACATCTGGAACAGCTCTTA | | 1268 | |
| QY | 241 | GlnAspGlySerLeuIleIleAspLeuGlyGlyAlaThrIleLeuTrpArgThrAlaAspGly | 260 | | |
| | | | | | |
| Db | 1270 | ----- | GGCCCAAGCTT----- | GACCAGACGCTGCGGTGCCGACAG----- | 1302 |
| QY | 261 | LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaIaIa | 280 | | |
| | | | | | |
| Db | 1303 | TTCTGTGTCGCCACCTGCTGAGCAGCATCTG----- | | 1333 | |
| QY | 281 | ArgProGlnCysProValGlyLeuAsnThrLeu | 291 | | |
| | | | | | |
| Db | 1333 | ---GCAAGTGTCTGTGTGAGAGTACTTACCTTCCTC | 1362 | | |

| | | | |
|----|---------|--|---------|
| Db | 3782406 | TTGTGGCGGAAATCAACGGGCTTACCATTTGAGAGAAATGAAACGGCGCTGGGTGGGTGGCT | 3782465 |
| QY | 28 | -----AanglyAlaLeuProAsn | 33 |
| Db | 3782466 | CAACGGATTGACTATCGACGACGCGGTTGCTTGAACACGACCAATTAGACGATTCCTTCC | 3782525 |
| QY | 34 | GlyAspArggGlyAlaArgGlySerArgPheAlaLeuTyrValArg-----ProLys | 50 |
| Db | 3782526 | GAGCTGGCGGCGGACGCTCGAGCTCGGTTCGAAAGCAAGACAGCACTAGTGGCTTCGA | 3782585 |
| QY | 51 | AlaAsnglyValLysProSerThrValHisValLleSerThrProGlnAlaSerLysAla | 70 |
| Db | 3782586 | CTACATGGCGCTTCCGCCCAAGGATATCCAGACGCGCGCTCCAGCCTTAAAGAGCCGGCG | 3782645 |
| QY | 71 | LleSerCysValgIyGln-----His-SerLleSerTyrThrLeuSerArgAsnGlnThr | 88 |
| Db | 3782646 | CGGTGGCTCAAGAGTCGGGTGTTCTTCACAAACACAGCCTGTGGTCAAGCGCGCCAGACCT | 3782705 |
| QY | 88 | rValValValGlnTyrThrHisAspLysAspThrAspMetPheGlnValGlnArgSerThr | 108 |
| Db | 3782706 | CGTTCAGTGCGCGCGCTGCGACGACCTCC----- | 3782733 |
| QY | 108 | rgLuserProLleAspPheValValThrAspThrLleSerGlySerGlnAsnThrAspGln | 128 |
| Db | 3782734 | -----GTCCAGCTCGGTGAGTTGCGAGATATCGAGATATCGAGCTCCAAAGCTTGACC | 3782783 |
| QY | 128 | uAlaGlnLleThrGlnSerThrLleSerArgPheAlaCysAsnArgLleValCysAspArgAs | 148 |
| Db | 3782784 | G----- | 3782844 |
| QY | 148 | ngLusProTyrThrAlaArgLlePheAlaAlaGlyPheAspSerSerLysAsnLlePheLe | 168 |
| Db | 3782785 | -CGGGCCATATACACGATCCGGAGAACCTCCGGGTGTCCACAACTCATGCTCCCTTCTT | 3782843 |
| QY | 168 | uGlyGlnLysAlaAlaLysTyrLysAsnProAspGlyHisMetAspArgLleuThrThrAs | 188 |
| Db | 3782844 | C---GACGACATCCGCCACCTGG-----GTTAGAGCGCGACGCTGCC | 3782882 |
| QY | 188 | ngLyVal---LeuValMetHisProArgGlyGlyPheThrGlnLysSerGlnProGlyVal | 207 |
| Db | 3782883 | GGGCAATGGCGGCGGATCCTGCGACGACCTTCCGCAACAACTCCGAC-----GGCGT | 3782933 |
| QY | 207 | lTyrArgGlnLleSerValCysGlyAspArgValTyrThrLeuArgGlyThrArgSerAlaGln | 227 |
| Db | 3782934 | CGTTCCGAGATTTGCGACGCTTGGAGGATCTGACGCG--CGCGGACCATGACCGGAACG | 3782992 |
| QY | 227 | ngLysArgGlyLysLeuValGlnSerGlnThrAsnValLeuGln----- | 241 |
| Db | 3782993 | TGACGAGCGGACATCCGCGACAGGACGCGCTCATTTGTTACTTCATCGACGACCGCA | 3783052 |
| QY | 242 | -----AspLysSerLeu-----LleAspLysCysGly----- | 250 |
| Db | 3783053 | GTGCATTTAGGATGGCAAGCTTTAACTACACCTGCGCGGACGCAAGTAAAGTCTTGGCAATCC | 3783112 |
| QY | 251 | -AlaThrLeuLeuThrPargThrAlaAspArgGlyLeuPheHisThrProThrGlnLysHisL | 270 |
| Db | 3783113 | ACAACGGGCGGTATGGCGGTTCCGA-----GTGTGCGATAGCCACCAACGCGCGGAC | 3783166 |
| QY | 270 | egLysAlaLeuArgGlnGlnLle-----AsnAlaAlaArgProGlnCys | 284 |
| Db | 3783167 | TGATCTGGCGGCGCAAGAG-ATGTGCGCGCTGCTGCTGCGCAATGCGCAGAGCTTGATG | 3783225 |
| QY | 284 | sProValGlyLeuAsnThrLeuAlaPhePro---SerLleAsnArgLysGlnValValGln | 303 |
| Db | 3783226 | C-----AAATAGCTGTCTTCCGACAGTGGCTGGCGCGGTGGAAGAAATCCCCAC | 3783273 |
| QY | 303 | uGlnLysGlnProThrAlaTyr-----LeuSerCysGlyHisVala | 316 |
| Db | 3783274 | GAGGCATCGGCGCTCTCGATTTAAAGATCAACCAACCAATGCGATCTGAGCATAGGAGATCA | 3783333 |
| QY | 316 | lHisGlyTyrHisAsnThrGlnHisArgSerAspThrGlnAlaAsnGlnArgLysCys-- | 335 |

```

Db      3783334 TCATAGTTAC-----TGCCAGACATATCAGATGGCAGTGTATTAATATGCCATGCG 3783387
Qy      336 -----PrometCysArgThrVal 341
Db      3783388 TGCCACTTATCCGCCAGCAGAACCGTC 3783415
RESULT 15
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 4.97e+04 Length: 4411529
Score: 106.00 Matches: 85
Percent Similarity: 35.90% Conservative: 55
Best Local Similarity: 21.79% Mismatches: 140
Query Match: 4.63% Indels: 112
DB: 3 Gaps: 19

US-10-041-030-4 (1-420) x US-09-103-840A-1 (1-4411529)
Qy      10 CysAlaProAsnLysGluProValLysTyrGlyGluLeuValLeuGlyTyr----- 27
Db      3790226 TGTGGCGCAATACACGGGCTTACATTGGAGAGATCGAACCGCTGGGGTGGTGGCT 3790285
Qy      28 -----AsnGlyAlaLeuProAsn 33
Db      3790286 CAACGATTGACTATGACGACGCGCTTCTGAAACCCAGCCATTAAGGAGTCTCTTC 3790345
Qy      34 GlyAspArgGlyArgArgLysSerArgPheAlaLeuTyrLysArg-----ProLys 50
Db      3790346 GAGCGTGGGGCGGAGCTGAGCTGCTTCAAGGCAAGACGAGCTAGTGGCGCTCGA 3790405
Qy      51 AlaAsnGlyValLysProSerThrValHisValHisSerThrProGlnAlaSerLysVala 70
Db      3790406 CTACATGGGCTTCGCGAGGATATCCAGACGCGCTTCAAGCCCTTGAAGAAGCCCGCG 3790465
Qy      71 HisSerCysLysGlyGln-----His-SerLysSerTyrThrLysSerArgAsnGlnH 88
Db      3790466 CCGTGCGCTCAAGGTGGTGTCTCAAAACAAACAGCTGTGTGTCAGCCCGCAGCGCT 3790525
Qy      88 rValValValGluTyrThrHisAspLysAspThrAspMetPheGlnValGlyArgSerTh 108
Db      3790526 CTTTCAGTGGCGCGCTCTGCAGACCTC----- 3790553
Qy      108 rGluSerProLLeaPhePheValValThrAspThrHisSerGlySerGlnAsnThrAspG 128
Db      3790554 -----GTTCGACGTGTGTGATTCGACGATGATCGCAGGCTGACCC 3790603
Qy      128 uAlaGlnIleThrGlnSerThrLysSerArgPheAlaCysArgIleValCysAspArgAs 148
Db      3790604 G----- 3790604
Qy      148 nGluProTyrThrAlaArgIlePheAlaAlaGlyPheAspSerSerLysAsnIlePheLe 168

```

```

Db      3790605 -CGGCGCTATCAAGGATCGCGGAAGCCCTCGCGCTCTCGACAAAGTCATGCTGTCTT 3790663
Qy      168 uGlyGluValAlaAlaLysTyrLysAsnProAspGlyHisMetAspGlyLeuThrThAs 188
Db      3790664 C---GACGACATTCGCGCACTG-----GTGAGGGCGGACGGTGGCC 3790702
Qy      188 nGlyVal---LeuValMetHisProArgGlyGlyPheThrGluGluSerGlnProGlyVa 207
Db      3790703 GGGCATGGCGCGGATACCTCGTGAACCGTTCGGACAAACTCGGCAC-----GGCGT 3790753
Qy      207 LTrpArgLuiLieserValCysGlyAspValTyrThrLeuArgGluThrArgSerAlaG 227
Db      3790754 CGTTGCGATTGTTCACACCTTGGAGCGATCTTGACCGG-GCGCGACCATGACCGAACG 3790812
Qy      227 nGlnArgGlyLysLeuValGluSerGluThrAsnValLeuGln----- 241
Db      3790813 TGACGAGCGGACATCGCGCAGAGGAGCGCTCATTTGTACTCTCATCGACCAAGCCGA 3790872
Qy      242 -----AspGlySerLeu-----IleAspLeuCysGly----- 250
Db      3790873 GTGCATTAAGATGGCAGCTTAACTACCGTCCGCGACAGTAAGTCTTGGCAATCC 3790932
Qy      251 -AlaThrLeuLeuTrpArgThrAlaAspGlyLeuPheHisThrProThrGlnLysHisI 270
Db      3790933 ACAACGGCGCTATGCGCGTTGCGA-----GTGTTCGATATGACCAACCGCGCGAC 3790986
Qy      270 eGluAlaLeuArgGlnGluIle-----AsnAlaAlaArgProGlnCy 284
Db      3790987 TGATCTGGCGCGACAGG-ATGTGCGCTGTGCTTCCATATGCGCCAGACTTGAAATG 3791045
Qy      284 sProValGlyLeuAsnThrLeuAlaPhePro---SerLleAsnArgLysGluValaValG 303
Db      3791046 C-----AATATGCTGTCTCTCCGAGTGGCTTGGCGCTGAAAAATCCGAC 3791093
Qy      303 uGluLysGlnProTrpAlaTyr-----LeuSerCysGlyHisVa 316
Db      3791094 GAGCCATGCGGCTCTGCGTATTAAGATCCACCCATCGCATCTGACACATAGGGCATCA 3791153
Qy      316 HisGlyTyrHisAsnTrpGlyHisArgSerAspThrGlnAlaAsnGluArgGluCys-- 335
Db      3791154 TCATAGTTAC-----TGCGACACATATCAGATGCGAGTGTATTAATATGCCATGCG 3791207
Qy      336 -----PrometCysArgThrVal 341
Db      3791208 TGCCACTTATCCGCCAGCAGAACCGTC 3791235

```

Search completed: December 9, 2004, 19:16:03
 Job time : 6915 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2004, 13:11:42 ; Search time 789 Seconds
(without alignments)
2794.369 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MFSPGQEHHCAPNKEPVKYG.....ATQLVGEQNCIKLIFGSPID 420

| Scoring table: | | |
|----------------|----------|-------------|
| | BLOSUM62 | |
| Xgapop | 10.0 | Xgapext 0.5 |
| Ygapop | 10.0 | Ygapext 0.5 |
| Fgapop | 6.0 | Fgapext 7.0 |
| Delop | 6.0 | Delext 7.0 |

Searched: 4134886 beqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```

Command line parameters: -DEV=xlp
-MODEL=frame+ p2n.model -O=59n2.1/USPTO.spool -OUT=10041030/runat 09122004 095516 7653/app.query.fasta_1.583
-DB=Genesecv 23Sep04 -P=FASTA+fasta -SUPRT=rrng -MINMATCH=0 -LOOPELT=0
-LOOPEXT=0 -UNITS=bits -START=1 END=-1 -MATRIX=blosum62 -TRANS=human0.0.cdi
-LIST=45 -DOCCALIGN=200 -GTRSCORE=pcr -THR.MAX=100 -THR.MIN=0 ALIGN=15
-MODE=LOCAL -OUTFMT=pfco -NORMscore -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
USER=UST10041030 @CGCN 1.1.885 @runat 09122004 095516.7653 -NCPU=6 -ICPU=3
-NO MMAP -LARGESERIES -NEG_SCORES=0 -WAIT -DSBPBLOCK=100 -LONLOG
-DEVEL.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

```
Database : N_Geneseq_23Sep04.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004s.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

| SUMMARIES | | | | | |
|------------|-------|-------------|--------|----|--------------------|
| Result No. | Score | Query Match | Length | DB | ID |
| 1 | 2390 | 100.0 | 1565 | 6 | AAL46323 Human M33 |
| 2 | 2280 | 100.0 | 1893 | 10 | ADQ30209 Human nov |
| 3 | 2280 | 100.0 | 4553 | 6 | ABQ78319 Nucleotid |
| 4 | 2280 | 100.0 | 3579 | 8 | ACC42349 Human MAP |
| 5 | 2283 | 99.7 | 1263 | 6 | ABLS8949 Human pel |
| 6 | 2283 | 99.7 | 1263 | 12 | ADP48670 Human pel |

| | | | | | | | |
|----|--------|------|-------|----|-----------|-----------|--------------|
| 7 | 2175.5 | 95.0 | 1260 | 6 | ABL584848 | ABL58448 | Murine pel |
| 8 | 2175.5 | 95.0 | 1260 | 12 | ADP486668 | ADP48666 | Murine pel |
| 9 | 2111.5 | 92.2 | 1717 | 8 | ACC42248 | ACC42248 | Mouse MAP |
| 10 | 1917 | 83.7 | 1257 | 6 | ABL58447 | ABL58447 | Human pel |
| 11 | 1917 | 83.7 | 1257 | 12 | ADP486666 | ADP48666 | Human pel |
| 12 | 1917 | 83.7 | 1304 | 8 | ABX05095 | ABX05095 | Human nov |
| 13 | 1917 | 83.7 | 3222 | 6 | AAI46320 | AAI46320 | Human M30 |
| 14 | 1917 | 83.7 | 3526 | 6 | AAI46319 | AAI46319 | Human M30 |
| 15 | 1917 | 83.7 | 7136 | 12 | ADP83102 | ADP83102 | Human pel |
| 16 | 1916 | 83.7 | 2736 | 6 | AAI46316 | AAI46316 | Rat M30 c |
| 17 | 1912 | 83.5 | 1257 | 6 | ABL58446 | ABL58446 | Murine pe |
| 18 | 1912 | 83.5 | 1257 | 12 | ADP486664 | ADP48664 | Murine pe |
| 19 | 1822 | 79.6 | 3846 | 6 | AAI46317 | AAI46317 | Human M30 |
| 20 | 1642 | 71.7 | 1338 | 6 | ABL58452 | ABL58452 | Human pel |
| 21 | 1642 | 71.7 | 1338 | 12 | ADP486674 | ADP486674 | Human pel |
| 22 | 1642 | 71.7 | 2683 | 10 | ADD71199 | ADD71199 | Human int |
| 23 | 1640 | 71.6 | 2589 | 6 | AAI46322 | AAI46322 | Human M31 |
| 24 | 1633 | 71.3 | 2588 | 10 | ADC30836 | ADC30836 | Human nov |
| 25 | 1632 | 71.3 | 1936 | 6 | AAI46321 | AAI46321 | Murine M3 |
| 26 | 1618 | 70.7 | 3254 | 6 | AAI46318 | AAI46318 | Human M30 |
| 27 | 1605 | 70.1 | 3332 | 3 | AAI66453 | AAI66453 | Human sec |
| 28 | 1574 | 68.7 | 2866 | 6 | AAH16312 | AAH16312 | Human sec |
| 29 | 1574 | 68.7 | 2866 | 4 | AAH78084 | AAH7808 | Nucleotic id |
| 30 | 1482.5 | 64.7 | 2508 | 10 | ADPB6804 | ADPB6804 | Human CDN |
| 31 | 1370 | 59.8 | 2183 | 4 | AAK94855 | AAK94855 | Human ful |
| 32 | 1370 | 59.8 | 2183 | 12 | ADL31993 | ADL31993 | Pull leng |
| 33 | 1302 | 56.9 | 2666 | 4 | ABL09073 | ABL09073 | Drosophili |
| 34 | 1291 | 56.4 | 16092 | 4 | AAH73420 | AAH73420 | Human imm |
| 35 | 1172 | 51.2 | 803 | 4 | AAH06690 | AAH06690 | Human CDN |
| 36 | 1102 | 48.1 | 592 | 10 | ADGC3115 | ADGC3115 | Human nov |
| 37 | 1075 | 46.9 | 723 | 5 | AA868681 | AA868681 | DNA encod |
| 38 | 1064 | 46.5 | 567 | 6 | AAI46314 | AAI46314 | Human M33 |
| 39 | 1050 | 45.9 | 570 | 6 | AAI46324 | AAI46324 | Human M32 |
| 40 | 1022 | 44.6 | 541 | 12 | ACH89787 | ACH89787 | Human gen |
| 41 | 996 | 42.5 | 31882 | 4 | ABL09072 | ABL09072 | Drosophili |
| 42 | 974 | 42.5 | 528 | 12 | ACH87309 | ACH87309 | Human gen |
| 43 | 904.5 | 39.5 | 869 | 10 | ADC32585 | ADC32585 | Human CDN |
| 44 | 858.5 | 37.5 | 741 | 4 | AAK93822 | AAK93822 | Human CDN |
| 45 | 858.5 | 37.5 | 741 | 4 | AAK932298 | AAK932298 | Human CDN |

ALIGNMENTS

```
RESULT 1
AAL46323
ID      AAL46323 standard, cDNA; 1565 BP
XX
AC      AAL46323;
```

KM Neurodegenerative disease; M30; M31; M32; M33; stroke;
KM Fragile X syndrome; Huntington's disease; Parkinson's disease;
KM Alzheimer's disease; multiple sclerosis; ovarian cancer;
KM Neurodegeneration; immune disorder; autoimmune disease; allergy;
KM infection; leukaemia, inflammation; neuroprotective; cerebroprotective;
KM immunosuppressive; cytostatic; neurotropic; antiparkinsonian; antiallergic
KM virucide; antiinflammatory; gene; ss.

OS Homo sapiens.

PN WO200221138-A2.

PD 14-MAR-2002

PF 07-SEP-2001; 2001WO-EP010366.

PR 07-SEP-2000; 2000US-00657479.

PA (AXAR-) AXARON BIOSCIENCE AG.

X

PI Schneider A, Hiemisch H, Rosener M, Klugmann M, Naim J;
PI Eisenhardt G, Kuner R, Lanahan A, Worley P, Spielvogel D, Scheek S;
DR WPI, 2002-292287/33.
DR P-PSDB; AAO17522.

PT Diagnosis of neurodegenerative disease comprises detecting level of M30-
XX family proteins.

PS Claim 3; Page 113-116; 130pp; German.

XX The present invention relates to a method of diagnosing neurodegenerative
CC diseases, comprising determining the concentration of a protein in a body
CC sample, where the protein may be M30 or a variant thereof, M31, M32 or
CC M33. The method is used to diagnose neurodegenerative diseases,
CC particularly stroke but also e.g. fragile X syndrome, Huntington's,
CC Parkinson's and Alzheimer's diseases, multiple sclerosis etc. Also
CC overexpression of M31 can be used for diagnosis of carcinoma and sarcoma,
CC especially ovarian cancer. The proteins can be used to identify specific
CC ligands, potentially useful for treating neurodegeneration, immune-system
CC disorders (e.g. autoimmune diseases, allergy, viral infection, leukemia,
CC inflammation etc.), carcinoma and sarcoma. Inhibitors of the interaction
CC between the proteins and the protein kinase IRAK-1 can be used to treat
CC neurodegeneration. The present sequence is a coding sequence of a protein
CC used in the method of the invention

XX Sequence 1565 BP; 393 A; 395 C; 441 G; 336 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6, 51e-200 Length: 1565
Score: 2290.00 Matches: 420
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Dbs: 6 Gaps: 0

US-10-041-030-4 (1-420) x AAL46323 (1-1565)

QY 1 MetPheSerProG1GInGluGluH1sCySA1AProAenLyAGluProVal1LyTyG1Y 20
DB 51 ATGTTTTCCTGGCGAGGAGAACTGCGCCCAATPAAGAGCACTGAATACGCG 110
QY 21 GluLeuVal1LeuG1YTyAenG1YAlaLeuProAenG1YAspArg1YArgArgLys 40
DB 111 GAGCTGGTGTCTCGGTACATAGTGTCTTACCAATGAGATAGAGAGAGAGAA 170
QY 41 SerA9PheAlaLeuTyTr1YsArgProLySA1aAenG1YVal1YProSerThrValHis 60
DB 171 AGTAGATTGGCCCTCAAGCGGCCCAAGCAATGTGTCAAACTCCGACCGTCCAT 230
QY 61 Val11SerThrProG1nAlaSer1YsAla11SerCValYAG1GInH1sSer11Ser 80
DB 231 GTGATATCCAGCCCGCCAGCATCCAAAGGCTATCAAGTGAAGGTCAACAGATATATCC 290
QY 81 TyTr1LeuSerA9sAenG1nThrVal1Val1G1uTyTrTh1sAsp1YsAspThraAP 100
DB 291 TACACTTGTCAAGAAATCAAGACTGTGTGTGAGATACACATATAGATACGAGT 350
QY 101 MetPheG1nVal1G1YArgSerThrG1uSerPro11LeAPheVal1ThraAspThr11e 120
DB 351 ATGTTTCAGGTGGCGAGATCAACAGAAACCTTATGACTGTGTGTCAACAGACGATT 410
QY 121 SerG1YserG1nAsThraSPG1u1aG1n11eThG1nSerThr11SerArgPheA1a 140
DB 411 TCTGGCAGCAAAACCGAGCAAGCCAGATCAACAGAGCCATATATCCAGGTTCGCG 470
QY 141 CySA9G1LeVal1CySA9PArgAenG1uProTyTrThra1aArg1ePheA1a1aG1YpHe 160
DB 471 TGCAGGATGTGTGGAGACAGAAATGAACTTACACAGCAGGATATTCGCCCGGATTT 530
QY 161 AspSerSer1YsAsn11PheLeuG1YG1uLySA1aAla1YTrP1YsAsnProAspG1Y 180
DB 531 GACTCTTCMAAAACATATATCTTGGAGAAAAGCGCAAAAGTGAAAAAACCCTCCGACGCG 590

QY 161 HisMetAspG1YLeuThrThraSPG1YVal1LeuVal1MetH1sProA9G1YG1YpHeThr 200
DB 591 CACATGATGGAGCTCACTAATATGAGCTCTGTGATGATCATCCAGAGGGGCTTCACC 650
QY 201 GluG1uSerG1nProG1YVal1TrPA9G1u11eSerVal1CySG1YAspVal1YTrThrLeu 220
DB 651 GAGGATCCCAAGCCCGGGGTCTGGCGGAGATCTGTCTGTGAGATGTGTACACTTG 710
QY 221 ArgG1uThraArgSerA1eG1nG1nArgG1YLyLeuVal1G1uSerG1nThraSnVal1Leu 240
DB 711 CGAGAAACCAAGTCCGCGCCACCAACAGAGAAAGCTGTGGAAGAGAACCAACCTCCG 770
QY 241 G1nAspG1YserLeu11AsPLeuCySG1YAlaThrLeuLeuTrPA9G1YThra1aSPG1Y 260
DB 771 CAGAGCGCTCCCTATTGACTGTGTGGGCCACTCTCTCTGAGAAACAGCAATGGG 830
QY 261 LeuPheH1sThrProThrG1nLyH1s11eG1uAlaLeuArgG1nG1u11eAsnA1a1a 280
DB 831 CTTTTCATDAGTCCAACTCAGAAAGCAGATAGAGCCCTCCGAGAGATTAACGCCGCC 890
QY 281 ArgProG1nCyProVal1G1YLeuAsnThrLeuA1aPheProSer11LeAsnArg1YSG1u 300
DB 891 CGGCTCAGTGTCTGTGGGCTCAACCCCTGGCTTCCCGACATCAACAGAAAGAG 950
QY 301 Val1Val1G1uG1uLyG1nProTrPA1aTyTrLeuSerCySG1YH1sVal1H1sG1YTrH1s 320
DB 951 GTGTGGAGAGAGACAGACCCCTGGGCAATATCTCACTTGTGGCCAGCTGACCGGATCCAC 1010
QY 321 AsnTrPA9G1YH1sArgSerAspThrG1uAlaAenG1uArgG1uCyProMetCyAArgThr 340
DB 1011 AACTGGGGCCATCCGAGATGACACGAGGCCAAAGAGAGAGTGTCCCATGTGCGAGACT 1070
QY 341 ValG1YProTyTrVal1ProLeuTrLeuG1YCySG1YVal1aG1YpHeTyTrVal1aSP1aG1Y 360
DB 1071 GTGGGCCCTTAGTGTCTCTGTGGCTGTGGTGGAGAGATTTATGTAGACCAAGAG 1130
QY 361 ProProThrH1sAlaPheThrProCySG1YH1sVal1CySerG1uLySerA1a1YsTyTr 380
DB 1131 CGGCCAATCATGCTTTTCACTCCCTGTGACACAGTGTCTCGAGAAAGTCTGCAGAAATAC 1190
QY 381 TrpSerG1n11eProLeuProH1sG1YThrH1sAlaPheH1sAla1aCySPProPheCyS 400
DB 1191 TGCTCTCAGATCCCGTGTCTCATATGAACTCATGATTTCAAGCTGTGACCTTTCTGT 1250
QY 401 AlaThrG1nLeuVal1G1YG1uG1nAsnCyS11eLyLeu11ePheG1G1YPro11eAsp 420
DB 1251 GCTACACAGCTGTGTGGGAGCAAACTGCATCAATTAATTTTCCAAGTCCAAATGAC 1310
RESULT 2
ADC30209
ID ADC30209 standard; cDNA, 1823 BP.
XX
AC ADC30209;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel cDNA sequence, SEQ ID NO:291.
XX
XX Human; diagnostic; drug screening; forensic; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; anti-Parkinsonian; nootropic;
KW neuroprotective; anti-nausea; antidiabetic; thrombolytic; vulnerary;
KW antitumor; osteopathic; immunosuppressive; anti-inflammatory; cytosclastic;
KW gene therapy; chromosome 14q21; gene; ss.
OS Homo sapiens.
XX
XX WO2003029271-A2.
XX
XX 10-APR-2003.
PD

XX 24-SEP-2002:~2002WC=US030474.
 XX 24-SEP-2001: 2004US-0324634P.
 XX (HYSE-) HYSEQ INC.
 XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T,
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Wang G,
 PI Haley-Vicente D, Drmanac RT,
 DR WPI: 2003-371981/35.
 DR P-PSDB; ADC31180.
 XX New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 PS claim 1; SEQ ID NO 291; 1185bp, English.
 XX The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1823 BP; 425 A; 483 C; 530 G; 385 T; 0 U; 0 Other:
 XX
 Alignment Scores:
 Pred. No.: 8.02e-200 Length: 1823
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-10-041-030-4 (1-420) x ADC30209 (1-1823)
 QY 1 MetPheSerProGlyValGlnGluGlnHisCysAlaProAsnLysGluProValLysTyrGly 20
 DB 265 ATGTTTTCCTCCCTGCGCAAGGAGAACTGCGCCCAATTAAGAGCGACAGGAATACGGG 324
 QY 21 GluLeuValValLeuGlyTyrAsnGlyValAlaLeuProAsnGlyAspArgGlyAspGly 40
 DB 325 GAGCTGTGTGTGCTCGGATCAATGCTTACCCCAATGAGATAGAGGACGAGGAAA 384

QY 41 SerArgPheAlaLeuTyrLysArgProLysValAlaAsnGlyValLysProSerThrValHis 60
 DB 385 AGTAAATTTTGCCTCTACAGAGCGGCCCAAGGCAAAATGGGTCAAAACCCGACCGCTCAT 444
 QY 61 ValLysSerThrProGlnAlaSerLysValAlaLysSerCysLysGlyGlnHisSerLysSer 80
 DB 445 GTGATATCCAGCCCGGAGCATCCAAAGGCTTACAGCTGCAGAGGTCAACACAGTATATCC 504
 QY 81 TyrTrpLysSerArgAsnGlnThrValValGlyTyrThrHisAspLysAspThrAsp 100
 DB 505 TACACTTGTCAAGGATCAGACTGTGGTGGTGAATACACACATGATAGATACCGAT 564
 QY 101 MetPheGlnValGlyArgSerThrGluSerProLysPheValValThrAspThrIle 120
 DB 565 ATGTTTACAGTGGGAGATCAACAAGCCCTTATGCACTTGGTGTCAACACAGATT 624
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 625 TCTGGACGCCAAGACAGGACGAGCCAGATCACAGAGCAACATATCCAGTTCCGCC 684
 QY 141 CysArgIleValCysAspArgAsnGluProTyrTrpAlaArgIlePheAlaAlaGlyPhe 160
 DB 685 TCGAGGATCGTGTGGACAGAGAAATGAACTTACAGACGCGATATTGCGCCCGGATTT 744
 QY 161 AspSerSerLysAsnLysPheLeuGlyGlyValAlaAlaLysTrpLysAspProAspGly 180
 DB 745 GACTCTTCCAAAACATATTTTGTGAGAAAGGACGAAAGTGGAAAAACCCCAACGCGC 804
 QY 181 HisMetAspGlyLysThrThrAsnGlyValLysValMetHisProArgGlyGlyPheThr 200
 DB 805 CACATGATGGCTCACTATATGAGCTCTGTGATGATCATCCAGAGGGGGCTTACCC 864
 QY 201 GluLysSerGlnProGlyValTrpArgLysLysSerValCysGlyAspValTyrThrLeu 220
 DB 865 GAGGATGCCAAGCCCGGGGTGTGGCGAGATCTGTCTGTGAGATGTGTACACTTG 924
 QY 221 ArgGluThrArgSerArgLysGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
 DB 925 CGAGAAACAGATCGGCGCCAGCAAGAGAAAGCTGTGGAAAGTGAGCAACGCTCTG 984
 QY 241 GlnAspGlySerLeuLysPheLysCysGlyValThrLeuThrArgThrAlaAspGly 260
 DB 985 CAGAGAGGCTCTCTCATATGACCTGTGTGGGCCACTCTCTTGGAGAACAGCAGATGGG 1044
 QY 261 LeuPheHisThrProThrGlnLysHisLysGlnAlaLeuArgGlnLysAlaAla 280
 DB 1045 CTTTTTCATATCCCAACTCAGAAAGCATAAGAAAGCCCTCGGACAGGATTAACCGCGCC 1104
 QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
 DB 1105 CCGCTCAGTGTCTCTGTGGGCTCAACACCTCGGCTTCCCGACATCAACGAGAAAGG 1164
 QY 301 ValValGlnGluLysGlnProThrAlaTyrLysSerCysGlyValHisValHisGlyTyrHis 320
 DB 1165 GTGTGAGAGAAAGAGCCCTGGGCAATCTTCAAGTTGTGGCCAGGTGACCGGATCCAC 1224
 QY 321 AsnTrpGlyHisArgSerAspThrGlnAlaAsnGluArgGluCysProMetCysArgThr 340
 DB 1225 AACTGGGGGCATCGAGGTACACGAGGCGCAAGAGGAGTGTCCCATGTGACAGACT 1284
 QY 341 ValGlyProTyrValProLeuThrLeuGlyCysGlnAlaGlyPheTyrValAspAlaGly 360
 DB 1285 GTGGGCGCCCTAATGTGCTCTGTGGCTGTGAGGAGAGATTATGTAAACGAGGA 1344
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
 DB 1345 CCGCCAACTCATGCTTTCATCTCTGTGACACGCTGTCTCGGAGAAAGTCTGCAAAATTC 1404
 QY 381 TrpSerGlnLysProLeuProHisGlyTyrHisValAsnHisAlaAlaCysProPheCys 400
 DB 1405 TGGTCAAGATCCCGTGTGCTCATGAGAACTTCAATTCACGCTGCTTCTGTCTGT 1464

QY 401 AAThGlnLeuValGlyGlnGluNaenCySileYLeuLeuIlePheGlnGlyProIleAsp 420
DB 1465 GCTACACAGCTGGTGGGAGCAAACTGCATCAATTAATTTTCCAAAGGTCATTGAC 1524

RESULT 3

ID ABQ78319 standard; DNA; 4563 BP.

AC ABQ78319;

DT 05-NOV-2002 (first entry)

DE Nucleotide sequence of human Pellino 2.

KM Pellino 2; cancer; Pellino 1; cancer treatment; epithelial cancer;

XX gastrointestinal tract cancer; gene; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 501..1763

FT /tag= a

FT /product= "Pellino 2"

FT 5'UTR 1764..4563

XX /tag= b

PN MO200259611-A2.

PD 01-AUG-2002.

XX 28-DEC-2001; 2001WO-US051368.

XX 02-JAN-2001; 2001US-0259502P.

XX (TUDA-) TUDARIK INC.

XX Powers S, Mu D, Xiang P, Peng Y;

XX MPI; 2002-619185/66.

XX P-PsDB; ABB78077.

XX Detecting cancer cells in mammalian sample, useful for identifying

XX inhibitors for treating cancer e.g. epithelial cancer, comprises

XX encoding, Pellino 1 and Pellino 2.

XX Claim 31; Page 57-61; 69pp; English.

XX The present sequence encodes human Pellino 2. The specification describes

XX a method for detecting cancer cells in biological sample from a mammal.

XX The method comprises detecting an overexpression of, or increase in copy

XX number of genes encoding, polypeptides Pellino 1 or Pellino 2. The method

XX is useful in detecting cancer or propensity to develop cancer, monitoring

XX the efficacy of cancer treatment, identifying inhibitors of Pellino 1 and

XX 2, inhibiting the expression and/or activity of Pellino 1 and 2 in cancer

XX cells, and treating cancer or inhibiting proliferation of cancer. The

XX cancer can be epithelial cancer, such as lung, colon, ovarian, breast,

XX prostate, kidney, stomach, bladder, or any cancer of the gastrointestinal

XX tract

SQ Sequence 4563 BP; 1146 A; 1028 C; 1106 G; 1283 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2 8e-199 Length: 4563

Score: 2290.00 Matches: 420

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-10-041-030-4 (1-420) X ABQ78319 (1-4563)

QY 1 MetPheSerProGlyGlnGluGluHisCySAlaProAsnLysGluProValLysTyrgly 20

DB 501 ATGTTTCCCTGGGCGCAGAGAAACATGCGGCCCCCAATTAAGACCGATGAATACGGG 560

QY 21 GIuLeuValIValLeuGlyTYrAsnGlyValAlaPProAsnGlyAspArgLys 40

DB 561 GAGCTGTGTGTCTGGGTACATATGTCCTTACCCTTACAGATAGAGAGCGAGGAAA 620

QY 41 SerArgPheAlaLeuTYrLysArgProLysAlaAsnGlyValLysProSerThrPAlHis 60

DB 621 AGTAGATTGGCTCTTACAGAGGCGCCAGGAATGGTGTCAAAACCCAGCACCGTTCAT 680

QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCySlyGlyGlnHisSerIleSer 80

DB 681 GTGATATCAGCGCCCGCAGGATCCAGGCTATCAGCTCAAAAGTCACACAGATATCC 740

QY 81 TYrThrLysSerArgAsnGlnThrValValValGlyTYrThrHisAspLysAspThrAsp 100

DB 741 TACACTTGTGTCAGGAATCAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 800

QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValThrAspThrIle 120

DB 801 ATGTTTCAGTGGGCGAGATCAACAGAAAGCCCTATCGACTTGTGTGTGTGTGTGTGT 860

QY 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140

DB 861 TCTGCGACCGAAGACAGGAGCGAAGCCAGATCAACAGAGACCATTCAGGTTCCGC 920

QY 141 CySArgIleValCySAspArgAsnGluProTYrThrAlaArgIlePheAlaIleGlyPhe 160

DB 921 TGCAGATCTGT 980

QY 161 AspSerSerLysAsnIlePheLeuGlyGlyLysAlaAlaTYrProLysAsnProAspGly 180

DB 981 GACTCTTCAAAATATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1040

QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgIleGlyPheThr 200

DB 1041 CACATGATGGGCTCATTACTATGATGCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1100

QY 201 GIuGluSerGlnProGlyValITPArgGluIleSerValCySArgLysAspValTYrThrLeu 220

DB 1101 GAGGAGTCCCGAGCCCGGGGTCTGGGCGAGATCTTGTGTGTGTGTGTGTGTGTGTGT 1160

QY 221 ArgGluThrArgSerAlaGlnGlnArgLysLysLeuValGluSerGlnThrAsnValLeu 240

DB 1161 CGAGAAACCAAGTCCGCCAGCAAGAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1220

QY 241 GlnAspGlySerLeuIleAspLeuCySArgLysValIleThrLeuLeuTYrArgThrAlaAspGly 260

DB 1221 CAGGACGGCTCCCTCATTTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1280

QY 261 LeuPheHisThrProThrGlnLysHisIleIleGlnAlaLeuArgGlnGlnIleAsnAlaIle 280

DB 1281 CTTTTCATCTCCCAATCAACAGCAATTAAGCCCTCGGCGAGAAATTAACCCGCC 1340

QY 281 ArgProGlnCySProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300

DB 1341 CGGCTCATGTCTCTGT 1400

QY 301 ValValGluGluLysGlnProTYrPAlaTYrLeuSerCySArgLysValIleGlyTYrHis 320

DB 1401 GTGTGTGAGAGAAAGCAAGCCCTGGGCGATATCTCAGTTGTGTGTGTGTGTGTGTGT 1460

QY 321 AsnTYrGlyHisArgSerAspThrGlnAlaAsnGlnValArgGluCySProMetCySArgThr 340

DB 1461 AACTGGGGCCATCGAGAGACAGAGAGCCCAAGAGAGAGTGTCCCATGTGTGTGTGTGT 1520

QY 341 ValGlyProTYrValProLeuTYrLeuGlyCySArgLysValIleGlyPheTYrValAspAlaGly 360

DB 1521 GTGGGCCCTTATGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1580

QY 361 ProProThrHisAlaPheThrProCySArgLysValIleCySArgLysSerAlaLysTYr 380

Db 1581 CCGCAACTGATGCTTACTCCCTGTGACACAGTGTGCTGGAGAAAGTCTGCAAAATAC 1640
 QY 381 TTPSGTGLIIEProleupProHISGlyThrHISAlaPheHISAlaAlaCysProPheCys 400
 Db 1641 TGGTTCAGATCCCGTGGCTCATGMAACTGATTCATTCACGCTGCTCCCTTCTGT 1700
 QY 401 AlaThrGluLeuValGlyGluGlnPheCysIleLeuLeuLeuPheGlnGlyProIleAsp 420
 Db 1701 GCTACACAGCTGGTGGGAGCAAACTGCATCAATTAATTTCCAAAGGTCCAAATTGAC 1760
 RESULT 4
 ACC42349
 ID ACC42349 standard; cDNA; 5579 BP.
 AC ACC42349;
 XX
 XX 22-MAY-2003 (first entry)
 DE Human MAP kinase cascade activator #59 cDNA.
 XX Human; Elki phosphorylation; Elki phosphorylation kinase; virucide;
 XX antiinflammatory; immunomodulator; cytosolic; antiallergic; anti-HIV;
 XX antihemetic; antidiabetic; antiaschemic; antiaschemic; gene therapy;
 XX inflammation; autoimmune disease; viral disease; cancer; diabetes;
 XX rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
 XX IGA nephritis; gene; ss.
 XX Homo sapiens.
 OS
 XX MO2003008589-A1.
 PN
 XX 30-JUN-2003.
 PD
 XX 15-JUL-2002; 2002MO-JP007174.
 PF
 XX 18-JUL-2001; 2001JP-00218204.
 PR 31-AUG-2001; 2001JP-00263450.
 PR 21-JUN-2002; 2002JP-00012176.
 XX
 XX (ASAH) ASAH KASEI KOGYO KK.
 PA
 XX Matsuzaki O, Matsumura A, Nagano Y, Suzuki N.
 PI
 XX WPI; 2003-229582/22.
 DR P-PSDB; ABR41083.
 XX
 PT Elki phosphorylation-associated gene and its encoded protein with MAP
 PT kinase cascade effect, applicable in diagnosis of and developing drugs
 PT for e.g. inflammation, autoimmune diseases, viral diseases and cancer.
 XX
 XX Claim 4; Page 631-640; 762pp; Japanese.
 PS
 XX The invention relates to a novel purified protein having Elki
 CC phosphorylation activity and/or an activity of activating Elki
 CC phosphorylation kinase. A protein of the invention has antiinflammatory,
 CC immunomodulator, virucide, cytosolic, antiallergic, antineumatic,
 CC antidiabetic, antiaschemic, and anti-HIV activity. The
 CC polynucleotides may have a use in gene therapy. The gene and its encoded
 CC protein are applicable in diagnosis of and developing drugs for e.g.
 CC inflammation, autoimmune diseases, viral diseases and cancer such as
 CC rheumatoid arthritis, diabetes, asthma, allergic rhinitis, AIDS, viral
 CC hepatitis and IGA nephritis. The present sequence is used in the
 CC exemplification of the invention
 CC
 SQ Sequence 5579 BP; 1535 A; 1111 C; 1214 G; 1719 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.69e-199
 Score: 2290.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 8
 Gaps: 0

Length: 5579
 Matches: 420
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-10-041-030-4 (1-420) x ACC42349 (1-5579)
 QY 1 MetPheSerProGlyGlnGluGlnHisCysAlaProAsnLeuGluProValIleuTyrgly 20
 Db 177 ATGTTTCCCTCCGCGCAGAGGAGAACCTCGCCGCCCAATTAAGAGGACATGAAATACGGG 236
 QY 21 GluLeuValIleuGlyTyrganglyAlaLeuProAsnGlyAspArgGlyArgGlyGly 40
 Db 237 GAGCTGGTGGTGTCTGGGTACATGATGCTTATCCCAATGAGATGAGAGAGCGAGAGAA 296
 QY 41 SerArgPheAlaLeuLeuTyrganglyProValIleuValIleuProSerThrValHis 60
 Db 297 AGTAGATTTGCTCTTCAAGAGCGGCCCAAGGCAATGATGTCAAACTCCAGACCGTCCAT 356
 QY 61 ValIleSerThrProGlnAlaSerIleuValIleuSerCysGlyGlyGlnHisSerIleu 80
 Db 357 GTGATATCCAGCGCCCGCAGGATCCCAAGCTATCAGCTGCAAGAGTCAACAGATATATCC 416
 QY 81 TyrThrLeuSerArgAngInThrValIleuValIleuTyrganglyHisAspIleuAsp 100
 Db 417 TACACTTGTCAAGAGAAATCAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 476
 QY 101 MetPheGlnValIleuArgSerThrGlySerProIleuAspPheValIleuAspThrIle 120
 Db 477 ATGTTTCAGGTGGGCGATCAACAGAAAGCCCTATGACTGTGTGTGTGTGTGTGTGTGTGT 536
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 537 TCTGGAGCGCAAGAACCGAGAGCCAGATCAGACAGACAGACAGATCCAGGTTCCGCG 596
 QY 141 CysArgIleValIleuAspArgAngInProTyrganglyHisAspIleuAspThrIle 160
 Db 597 TGCAGGATGT 656
 QY 161 AspSerSerIleuAsnIleuPheLeuGlyGlyValIleuValIleuTyrganglyHisAsp 180
 Db 657 GACTCTTCGAAAGAAATATTTCTTGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 716
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleuValIleuValIleuValIleuValIleu 200
 Db 717 CACATGATGGGCTCATCTAATGCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 776
 QY 201 GluGluSerGlnProGlyValIleuValIleuValIleuValIleuValIleuValIleu 220
 Db 777 GAGGAGTCCAGCCCGGCTGT 836
 QY 221 ArgGluThrArgSerAlaGlnIleuValIleuValIleuValIleuValIleuValIleu 240
 Db 837 CGAGAAACAGAGTGGGCGCAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 896
 QY 241 GlnAspGlySerIleuLeuAspLeuCysGlyAlaThrLeuLeuThrArgThrAlaAspGly 260
 Db 897 CAGGAGGCTCTCTTATTAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 956
 QY 261 LeuPheHisThrProThrGlnIleuHisIleuValIleuValIleuValIleuValIleu 280
 Db 957 CTTTTTCATATCCCAATCAGAGACATAGAGAGCCCTCCGAGAGAGATTTAAACCGCCGC 1016
 QY 281 ArgProGlnCysProAlaIleuLeuAsnThrAlaPheProSerIleuAsnArgIleuGlu 300
 Db 1017 CGGCTCAATGTCTGT 1076
 QY 301 ValIleuGluGluGlnProThrAlaTyrganglyCysGlyHisIleuValIleuValIleu 320
 Db 1077 GT 1136
 QY 321 AsnThrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
 Db 1137 AACTGGGGGCATCGGATGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1196
 QY 341 ValGluProTyrganglyProleuThrPheGluGlyCysGluAlaGlyPheTyrgangly 360

```

Db      1197  GTGGCCCCCTATATGTCCTCTGCTGTGCTGTGAGGACAGATTATATGACGACGAGA 1256
QY      361  PropiothrinAlaAphethrProCysGlyHiseValCysSerGlnLysSerAlaLysTyr 380
Db      1257  CCGCCAACTCAATGCTTCACTCCCTGTGGACAGTGTGCTCGAGAAAGTCTCAAAATTC 1316
QY      381  TTPSerGlnIleProLeuProHiseGlyThrHiseAlaAphethrAlaAlaCysProPheCys 400
Db      1317  TGGTCTCAGATCCCGTGTGCTCATGAACTCATGATTCACGCTGCTGCTTCTCT 1376
QY      401  AlathrGlnLeuValGlyGlnGlnAenCysIleLysLeuIlePheGlnGlyProIleAsp 420
Db      1377  GCTACACAGCTGTGGGGAGCAAACTCATCAATTAATTTCCAAAGTCCAAATTGAC 1436

RESULT 5
ABLS8449
ID      ABLS8449 standard; DNA; 1263 BP.
XX
AC      ABLS8449;
XX
DT      30-JUL-2002 (first entry)
XX
DE      Human pellino-2 polypeptide coding sequence.
XX
KW      Nuclear factor-kappaB; NF-kappaB; p38; pellino; interleukin-1; virucide;
KW      antibacterial; fungicide; protozoacide; antilashmatic; antirheumatic;
KW      antiarthritic; antiinflammatory; antithrombotic; neuroprotective;
KW      nocotropic; antitumor; human; pellino-2; gene; de.
XX
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
FT      CDS      1..1263
FT      FT      /tag=a
FT      FT      /product="pellino-2"
XX
PD      WO200183739-A2.
XX
PD      08-NOV-2001.
XX
PF      27-APR-2001; 2001WO-US013676.
XX
PR      28-APR-2000; 2000US-0200198P.
XX
PA      (IMMUNEX CORP.
XX
PI      Bird TA, Cosman DJ;
XX
XX      WPI; 2002-066532/09.
XX      P-PSDB; ABB07922.
XX
PT      New Pellino polypeptides for identifying compounds that alter polypeptide
PT      activity, treating pathogenic infection or inhibiting apoptosis, are
PT      capable of stimulating nuclear factor-kappaB- or p38-dependent
PT      transcription.
XX
XX      Claim 3; Page 61; 70pp; English.
XX
XX      The invention provides polypeptides capable of stimulating nuclear factor
XX      (NF)-kappaB-dependent transcription or p38-dependent transcription,
XX      referred as Pellino polypeptides. The pellino polypeptides are useful for
XX      identifying modulators that alter the pellino polypeptide and pellino
XX      dominant-negative activity. They are also useful for identifying
XX      compounds that inhibit the binding activity of the polypeptides and to
XX      study cell-signal transduction. They are useful for preventing or
XX      treating infection by a pathogen such as virus, bacterial, fungi, algae
XX      or protozoa, or inhibiting apoptosis. Dominant-negative pellino
XX      polypeptides are useful for treating inflammatory conditions such as
XX      asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's
XX      disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and
XX      also for inhibiting mitogen activated protein (MAP) kinase-activated
XX      pathways. Pellino polypeptides and polynucleotides are useful to identify
XX      small molecule inhibitors of protein association or function of Pellino,

```

CC and other molecules involved in interleukin (IL)-1 signaling. The present sequence represents a human pellino-2 polypeptide coding sequence

Sequence 1263 BP, 317 A, 336 C, 351 G, 259 T, 0 U, 0 Other;

Alignment Scores:

| Pred. No.: | 2,13e-199 | Length: | 1263 |
|------------------------|-----------|---------------|------|
| Score: | 2283.00 | Matches: | 419 |
| Percent Similarity: | 99.76% | Conservative: | 0 |
| Best Local Similarity: | 99.76% | Mismatches: | 1 |
| Query Match: | 99.69% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |

US-10-041-030-4 (1-420) x ABLS8449 (1-1263)

```

QY      1  MetPheSerProGlyGlnGlnGlnHiseCysAlaProAenLysGlnProValLysTyrGly 20
Db      1  ATGTTTCCCTGGCCAGAGAGAAACATCGCCCCCAATTAAGAGCCAGTGAAATACGGG 60
QY      21  GlnLeuValValLeuGlyTyrAsnGlyValAlaLeuProAenGlyAspArgGlyArgArgLys 40
Db      61  GAGCTGGTGTGCTCGGTTACATGCTCTTTACCAATGAGATAGAGACGAGAGAAA 120
QY      41  SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db      121  AGTACATTTCCTCTTACAGCGGGCCAAAGCAATGTGTCAAAACCAGACCGTCCAT 180
QY      61  ValIleSerThrProGlnAlaSerLysValAlaIleSerCysLysGlyGlnHiseSerIleSer 80
Db      181  GTGATATCCAGGCCCCAGGACATCCAAAGGCTATCAGCTGCAAGGTCAACACATATATCC 240
QY      81  TyrThrLysSerArgAsnGlnThrValValValGlyTyrThrHiseAspLysAspThrAsp 100
Db      241  TACACTTTGTCAAGAAATCAGACTGTGTGTGTGAGTACACATGATTAAGATTCGAT 300
QY      101  MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValValThrAspThrIle 120
Db      301  ATGTTTCAGTGGGAGATCAACAGAAAGCCCTATGACTTCGTTGCACAGACACGATT 360
QY      121  SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      361  TCTGGACGCCAGAACACGACGAGAGCCCAAGATCACACAGACACCAATATCCAGGTGCC 420
QY      141  CysArgIleValCysAspArgAsnGlnProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db      421  TCGAGGATCGTGTGGAAGAAATGAACCTTACACAGACCGAATTTGCCGCCGATTT 480
QY      161  AspSerSerLysAsnIlePheLeuGlyGlnLysAlaAlaLysTyrLysAsnProAspGly 180
Db      481  GACTCTTCAAAAACATATATTTCTTGAGATAAAGCAGCAAGTGGAATAAACCCGACGCG 540
QY      181  HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db      541  CACATGATGGAGCTCACTACTAATAGGCTCTGAGTGAATCCACGAGGGGGCTTCAAC 600
QY      201  GlnGlnSerGlnProGlyValTyrArgGlnLysSerValCysGlyAspValTyrThrLeu 220
Db      601  GAGAGTCCCAAGCCCGGGGTCTGGCGCAGATCTGTCTGTGAGATGTATACACTTG 660
QY      221  ArgGlnThrArgSerAlaGlnGlnArgGlyLysLeuValGlnSerGlnThrAsnValLeu 240
Db      661  CGAGAAACCAAGTGGCGGCCCAACCAAGAGAAAGCTGGTGAAGAGAACCAACACTCTG 720
QY      241  GlnAspGlySerLeuIleAspLeuCysGlyValThrLeuLeuTyrArgThrAlaAspGly 260
Db      721  CAGAGCGGCTCCCTATTGACCTGTGTGGGCACTCTCTCTGAGAAACAGCAATGAG 780
QY      261  LeuPheHisThrProThrGlnLysHisIleGlnAlaLeuArgGlnGlnLysAlaAla 280
Db      781  CTTTTCATCTCACTCAACATCAGACACATTAAGCCCTCGGCGAGGATTAAGCCGCC 840
QY      281  ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGln 300

```


Sequence 1260 BP; 309 A; 357 C; 359 G; 235 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | 1,52e-189 | Length: | 1260 |
|------------------------|-----------|---------------|------|
| Score: | 2175.50 | Matches: | 399 |
| Percent Similarity: | 97.38% | Conservative: | 10 |
| Best Local Similarity: | 95.00% | Mismatches: | 10 |
| Query Match: | 95.00% | Indels: | 1 |
| DB: | 12 | Gaps: | 1 |

US-10-041-030-4 (1-420) x ADP48668 (1-1260)

```

QY 1 MetPheSerProG1yGlnGluGlnHisCysAlaProAlaSerGluProValIleYrGly 20
Db 1 ATGTTTTCCTCCGGCCAGAGAGAACCCACCCCAACAAGAGCGGTGAATACGGG 60
QY 21 GluLeuValValIleuGlyTyrAsnGlyAlaLeuProAlaSerGlyValIleYrGly 40
Db 61 GAGCTGGTGTCTCTGGGCTTACAAATGCTTAACTTAATGCTGACAGGGGAGAGAA 120
QY 41 SerArgPheAlaLeuTyrIleYrAsnGlyProGlyAlaSerGlyValIleYrGly 60
Db 121 AGGAGATTGCTCTTAAAGCGGACCTTAAGCGGACCTTAAGCGGACCTTAAGCGG 180
QY 61 ValIleSerThrProGlnAlaSerGlyValIleSerCysAlaSerGlyGlnHisSerIle 80
Db 181 ATGCTCTCACACACAGGCGCTTCAAGGCTCAAGCTCAAGGCTCAAGGCTCAAGG 240
QY 81 TyrThrLeuSerArgAsnGlnThrValValIleGlyTyrThrHisAspIleYrAsp 100
Db 241 TACACGTTGTCAAGGAGCAGACGAGTGTGTGAGTATACACACGATTAAGACAGG 300
QY 101 MetPheGlnValIleYrSerThrGlnSerProIleAspPheValIleThrAspThr 120
Db 301 ATGTTTCAAGTGTGGCGGTCAACAGAAACCCCATTTGCTGTGTCAACAGACAGG 360
QY 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPhe 140
Db 361 TCCGGCGGTCAAGACGAAAT---GCCGAGATCACACAGACCACTCTCTGTTCCGA 417
QY 141 CysArgIleValIleCysAspArgAsnGluProTyrThrAlaArgIlePheAla 160
Db 418 TCCAGGATCGTGTGTACAGAAACGAGCATATACAGACGATATTCCTGGCAGGAT 477
QY 161 AspSerSerIleAsnIlePheLeuGlyValIleValIleValIleValIleVal 180
Db 478 GATTTCTCAAAAATATCTTTCTTGAGAGAAAGCAGAAATGAAAAACCTGATGGA 537
QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleValIleValIleValIleVal 200
Db 538 CACATGAGTGAATCTCACTCAATGCTGTCTTAATGATGACCCCGAAGAGGCTTCA 597
QY 201 GluGluSerGlnProGlyValIleTyrArgGluIleSerValIleCysGlyAspVal 220
Db 598 GAGGATCCAGCTGAGCTGAGAGAGATCTCTGTGTGTGGGATGTGAACCTTG 657
QY 221 ArgGluThrArgSerAlaGlnGlnIleValIleValIleValIleValIleVal 240
Db 658 CAGAGACCAAGGTCGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717
QY 241 GlnAspGlySerLeuIleAspLeuValIleValIleValIleValIleValIle 260
Db 718 CAAGACGGCTCTCTTATGACCTGTGTGGGCACTCTCTCTGAGGAACCCGAGATG 777
QY 261 LeuPheHisThrProThrGlnIleValIleValIleValIleValIleValIle 280
Db 778 CTTTTCACGCTCTCTCTCAAGAGCAGATAGAGAGAGAGAGAGAGAGAGAGAGAG 837
QY 281 ArgProGlnCysProValIleValIleValIleValIleValIleValIleVal 300
Db 838 CGAGCCAGAGGCGGCTGTGGCTTAAACCCCTGTCCCTCCAGCATCAACCGAGAG 897
QY 301 ValIleGluGluValIleProTyrIleValIleValIleValIleValIleVal 320

```

```

Db 898 GTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957
QY 321 AsnTyrGlyHisArgSerAspThrGlnAlaSerGluValIleValIleValIle 340
Db 958 AGCTGGGGCCATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1017
QY 341 ValGlyProTyrValIleProLeuTyrLeuGlyCysGlyValIleValIleVal 360
Db 1018 GTGGGCTCTTACGCTCTCTCTGTGGCTGTGAGAGAGAGAGAGAGAGAGAGAG 1077
QY 361 ProProThrHisAlaPheThrProCysGlyHisValIleCysSerGluValIle 380
Db 1078 CCCCCAATCAAGCTTTCAACCCCTGCGGAGCAGCTGTTCAGAAAGCTTCAGAGTAC 1137
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysPro 400
Db 1138 TGTGTGAGATTCCTCCATCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1197
QY 401 AlaThrGlnLeuValIleGluGlnAsnCysIleLeuValIlePheGlnGlyPro 420
Db 1198 GCCAGCAGCTGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257

RESULT 9
ACC42348
ID ACC42348 standard; cDNA; 1717 BP.
XX
AC ACC42348;
XX
DT 22-MAY-2003 (first entry)
XX
DE Mouse MAP kinase cascade activator #9 cDNA.
XX
KW Mouse, Etk1 phosphorylation; Etk1 phosphorylation kinase; virucide;
XX antiinflammatory; immunomodulator; cytotoxic; antiallergic; anti-HIV;
XX antirheumatic; antidiabetic; antidiabetic; antidiabetic; gene therapy;
XX inflammation; autoimmune disease; viral disease; cancer; diabetes;
XX rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
XX IGA nephritis; gene; ss.
XX
OS Mus musculus.
XX
PN W02003008589-A1.
XX
PD 30-JAN-2003.
XX
PF 15-JUL-2002; 2002MO-JP007174.
XX
PR 18-JUL-2001; 2001JP-00218204.
PR 31-AUG-2001; 2001JP-00263450.
PR 21-JAN-2002; 2002JP-00012176.
XX
PA (ASAH) ASAH KASEI KOGYO KK.
PI
PI Matsuzaki O, Matsuda A, Nagano Y, Suzuki N;
XX
DR MPI: 2003-229582/22.
DR P-PSDB; ABR41082.
XX
PT Etk1 phosphorylation-associated gene and its encoded protein with MAP
PT kinase cascade effect, applicable in diagnosis of and developing drugs
PT for e.g. inflammations, autoimmune diseases, viral diseases and cancer.
XX
PS Claim 4; Page 623-627; 762pp; Japanese.
XX
XX The invention relates to a novel purified protein having Etk1
XX phosphorylation activity and/or an activity of activating Etk1
XX phosphorylation kinase. A protein of the invention has antiinflammatory,
XX immunomodulator, virucide, cytotoxic, antiallergic, antirheumatic,
XX antidiabetic, antidiabetic, antidiabetic, and anti-HIV activity. The
XX polynucleotides may have a use in gene therapy. The gene and its encoded
XX protein are applicable in diagnosis of and developing drugs for e.g.
XX inflammations, autoimmune diseases, viral diseases and cancer such as

```


XX 11-DEC-2002; 2002US-00317250.
 PR (IMV) IMMUNEX CORP.
 PA (CLEV-) CLEVELAND CLINIC FOUND.
 XX Bird TA, Cosman DJ, Li X;
 XX WPI; 2004-480927/45.
 DR P-PSDB; ADP48667.
 XX
 PT Identifying inhibitors of IL-1 signaling, useful for treating e.g.,
 PT asthma, and rheumatoid arthritis, comprises assaying the association of a
 PT Pellino-1 polypeptide with IRAK-4 in the presence of a test compound.
 XX
 XX Claim 9; SEQ ID NO 3; 81pp; English.
 CC The present invention describes a method for identifying compounds (C)
 CC that inhibit nuclear factor kappa B (NF-kB)-dependent transcription or
 CC p38-dependent transcription. The method comprises: (1) mixing a test
 CC compound with a Pellino-1 polypeptide; (b) assaying the association of
 CC the Pellino-1 polypeptide with IL-1 receptor-associated kinase 4 (IRAK-4)
 CC in the presence of the test compound; and (c) determining whether the
 CC test compound inhibits the association of the Pellino-1 polypeptide with
 CC a binding partner. Also described: (1) an inhibitory nucleic acid that
 CC binds to a nucleic acid encoding an amino acid sequence as described
 CC above, where the presence of the inhibitory nucleic acid within a cell
 CC inhibits the association of Pellino-1 with IRAK-4; and (2) an inhibitory
 CC polypeptide comprising an antibody fragment that binds to a polypeptide
 CC comprising an amino acid sequence as described above, where the presence
 CC of the inhibitory polypeptide within a cell inhibits the association of
 CC Pellino-1 with IRAK-4. (C) have antiartherosclerotic, antiarthritic,
 CC antiasthmatic, antiinflammatory, antirheumatic, antiulcer,
 CC gastrointestinal, neuroprotective and nociceptive activities, and can be
 CC used as an IL-1 antagonist, IRAK protein kinase family inhibitor, and NF
 CC kappa B inducing kinase inhibitor. The methods and compositions of the
 CC present invention are useful for the prevention and/or treatment of
 CC diseases or conditions associated with aberrant expression or activity of
 CC the Pellino-1 polypeptide, such as asthma, rheumatoid arthritis,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC atherosclerosis and Alzheimer's disease. The present sequence encodes
 CC human Pellino-1, which is used in the exemplification of the present
 CC invention. The human Pellino-1 gene is located on chromosome 2, more
 CC specifically to 2p13.3.
 CC
 XX
 SQ Sequence 1257 BP; 368 A; 264 C; 296 G; 329 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 7,49e-166 Length: 1257
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.48% Conservative: 38
 Best Local Similarity: 81.43% Mismatches: 38
 Query Match: 83.71% Indels: 2
 DB: 12 Gaps: 1
 US-10-041-030-4 (1-420) x ADP48666 (1-1257)
 QY 1 MetPheSerProGlyInGluGluIuH1sCyAlaProAbnlyGluProVallyIysTYGly 20
 DB 1 ATGTTTCTCTGATCAAGAAATCAT-----CAATCTAAAGCACCAAGTAAATATGTG 54
 QY 21 GluLeuValAlaLeuGlyTYaenGlyAlaLeuProAbnlyGlyAspArgGlyArgGly 40
 DB 55 GAACATATGCTTAGATTAATATGATCTCTCCAAACCGTATGAGAGAGAGAGAA 114
 QY 41 SerArgPheAlaLeuTYrlyAspArgProlyAlaAsnGlyVallylProSerThrValHis 60
 DB 115 AGTAGGTTTGCTTGTATTAAGACCTTAAGGCAAAATGGGTGAAGCCCACTGTCAT 174
 QY 61 ValIleSerThrProGlnAlaSerlyAlaIleSerCylsGlyGlnIHisSerIleSer 80
 DB 175 ATTGCTTGACTCTCTCAAGGCTGCAAGGCAATTAAGCAAAACACAGATACATATCA 234

QY 81 TyrThrLeuSerArgAbnGlnThrValAlaGluTYrThrHisAspIysAspThrAsp 100
 DB 235 TATACCTTATCTCGGGCCAGACTGTGTGTGTAATATATCATGACAGCAACACCGAT 294
 QY 101 MetPheGlnValAlaIysSerThrGluSerProIleAspPheValAlaThrAspThrIle 120
 DB 295 ATGTTTCAAGATTGGCCGGTGTGACTGAAAGCCCATTTGATTTGTATTACTGACAGGTT 354
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 355 CCGAAGATCAAGAAATTAATCTGTATACAGTACAGTACAAAGACATATCAAGATTGGCC 414
 QY 141 CysArgIleValAlaCysAspArgAbnGluProTYrThrAlaArgIlePheAlaAlaGlyPhe 160
 DB 415 TGCAGATCATATGTGAACGGAATCTCTCCCTTACGACGACGATTTATCTCGAGGCTT 474
 QY 161 AspSerSerIysAsnIlePheLeuGlyGluValAlaAlaIleTYrTrpIysAsnProAspGly 180
 DB 475 GACTCATCAAAAACATCTTTCTTGGGAGAAAGCTGCCAAATGGAAACATCAGATTGA 534
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisAspArgGlyIysPheThr 200
 DB 535 CAGATGATGGCTTGACCACTAATGATGTTCTTTGTATGATCATCAGCAATGGGTTCA 594
 QY 201 GluGluSerGlnProGlyValAlaTrpArgGluIleSerValCysGlyAspValTYrThrLeu 220
 DB 595 GAAGACTCCAAAGCTGGAAATATGAGAAATATCGCTGTGGAAATGATATTACCTTA 654
 QY 221 ArgGluThrArgSerAlaGlnGlnIleArgGlyValIleValGluSerGlnThrAsnValLeu 240
 DB 655 GTGAAACCAAGATCGCTCAAGCAAGAAAGAAATGTGAAATTAACCAATCACTGTTA 714
 QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTrpArgThrAlaAspGly 260
 DB 715 CAAGATGGCTCGTTATTAATGACCTGTGTGCAACATTGTTATGGCTATGCAAGAGG 774
 QY 261 LeuPheHisThrProThrGlnIlyHisIleGluAlaLeuArgGlnGluIleAsnAlaAla 280
 DB 775 CTTTCCCAACATCTTCACTGAAAGCAATTTAAGACTTTAAGACAGAAATCAATGACGCA 834
 QY 281 ArgProGlnCysProValGlyIleAsnThrLeuAlaPheProSerIleAsnArgIlyGlu 300
 DB 835 CAGCTCAGTGGCTGTAGGCTTCAACACATGACATTTCTTAGTATGAAGAGAAAGAC 864
 QY 301 ValAlaGluGluIysGlnProTrpAlaTYrLeuSerCysGlyHisValHisGlyTYrHis 320
 DB 895 GTTGATGATGAAGAAACAAACATGCGGTATATCTAACTGGCGCATGATGCTATCAT 954
 QY 321 AsnTrpGlyHisIleArgSerAspThrGluAlaAsnGluArgGluCysProMetCysAlaGly 340
 DB 955 AACTGGGGAAACAAAGAAACGATGTCAAAGATCGGAATGTCCTTAGTGTAGTCT 1014
 QY 341 ValGlyProTYrValProLeuTrpLeuGlyCysGluAlaGlyPheTYrValAspAlaGly 360
 DB 1015 GTTGTCTCTATGTTCTCTGTGCTGTGATGTGAAGCTGAAATTTATGTGAGCCCGG 1074
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlnIlySerAlaIlyTYr 380
 DB 1075 CTTCCAAACCAATCGTTTACCGCGGTGGCAATGTGTTTCAAGAAAGCAACTGCTAT 1134
 QY 381 TrpSerGlnIleProLeuProHisGlyTYrHisAlaPheHisAlaAlaCysProPheCys 400
 DB 1135 TGGTCCCAATCCCACTTCTCATGATGATCTCACTTTTCATGACGCGCTGCTTGTGT 1194
 QY 401 AlaThrGlnLeuValAlaGlyGluIlyAsnCysIleIysLeuIlePheGlnGlyProIleAsp 420
 DB 1195 GCAATCAATGCTGCTGTGAACAAGCTACATCAAGCTTATTTTCAAGAACTCTAGAC 1254
 RESULT 12
 ID ABX05095
 XX ABX05095 standard; cDNA; 1304 BP.
 AC ABX05095;

DB 1178 CCTCCACCCATGCGTTTACGCCGTGTGGCATGTGTGTTTCAGAAAAGCAACTGCTAT 1237
 QY 381 TTPSerglnIleProleuProHiseGlyThrHisAlaPheHisAlaIleCysProPheCys 400
 DB 1238 TGGTCCCAATCCCACTTCCTCATGTGACTCATCTTTCATGACACCTGCTCTTTTGT 1297
 QY 401 AlaThrGluLeuValGlyGluGlnAenCysIleLeuLeuIlePheGlnGlyProIleAsp 420
 DB 1298 GCACATCACTTGGCTGTGTGAACAGGCTACATCAAGCTTATTTTTCAGAGACCTTAGAC 1357

RESULT 14
 ID AAL46319
 AAL46319 standard; cDNA; 3526 BP.
 XX AAL46319;
 DT 19-JUL-2002 (first entry)
 DE Human M30 variant C coding sequence SEQ ID NO: 7.
 XX
 KM Neurodegenerative disease; M30; M31; M32; M33; stroke;
 KM fragile X syndrome; Huntington's disease; Parkinson's disease;
 KM Alzheimer's disease; multiple sclerosis; ovarian cancer;
 KM neurodegeneration; immune disorder; autoimmune disease; allergy;
 KM infection; leukaemia; inflammation; neuroprotective; cerebroprotective;
 KM immunosuppressive; cytostatic; nootropic; antiparkinsonian; anti-allergic;
 KM virucide; antiinflammatory; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200221138-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 07-SEP-2001; 2001WO-EP010366.
 XX
 PR 07-SEP-2000; 2000US-00657479.
 XX
 PA (AXAR-) AXARON BIOSCIENCE AG.
 XX
 PI Schneider A, Hiemisch H, Rossner M, Klugmann M, Naim J,
 PI Eisenhardt G, Kuner R, Lanhahn A, Worley P, Spielvogel D, Scheek S;
 DR WPI; 2002-292287/33.
 DR P-PSDB; AAO17518.
 PT
 PT Diagnosis of neurodegenerative disease comprises detecting level of M30-family proteins.
 PS
 PS Claim 3; Page 91-95; 130pp; German.
 CC The present invention relates to a method of diagnosing neurodegenerative
 CC diseases, comprising determining the concentration of a protein in a body
 CC sample, where the protein may be M30 or a variant thereof, M31, M32 or
 CC M33. The method is used to diagnose neurodegenerative diseases,
 CC particularly stroke but also e.g. fragile X syndrome, Huntington's,
 CC Parkinson's and Alzheimer's diseases, multiple sclerosis etc. Also
 CC overexpression of M31 can be used for diagnosis of carcinoma and sarcoma,
 CC especially ovarian cancer. The proteins can be used to identify specific
 CC ligands, potentially useful for treating neurodegeneration, immune-system
 CC disorders (e.g. autoimmune diseases, allergy, viral infection, leukaemia,
 CC inflammation etc.), carcinoma and sarcoma. Inhibitors of the interaction
 CC between the proteins and the protein kinase IRAK-1 can be used to treat
 CC neurodegeneration. The present sequence is a coding sequence of a protein
 CC used in the method of the invention

XX
 SQ Sequence 3526 BP; 1076 A; 618 C; 605 G; 1147 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3 06e-165 Length: 3526
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.48% Mismatches: 38
 Best Local Similarity: 81.43%

M31, M32, M33

Query Match: 83.71% Indels: 2
 DB: 6 Gaps: 1
 US-10-041-030-4 (1-420) x AAL46319 (1-3526)

QY 1 MetPheSerProGlyGlnGluGlnHisCysAlaProAenGlyGluProValIleSerGly 20
 DB 408 ATGTTTCTCCGATCAAGAAATATCAT-----CCATCTAAAGCAGCATTAATATGCT 461
 QY 21 GluLeuValIleLeuGlyIleThrAsnGlyIleLeuProAsnGlyAspArgGlyAsp 40
 DB 462 GAACATCATGTCTTGGGTATATATGGTCTCTCCCAATATGGCATATGAGAAAGGAGAA 521
 QY 41 SerArgPheAlaLeuThrIleValArgProValIleAsnGlyValIlePheSerThrValHis 60
 DB 522 AGTAGGTTTGCTTGTCTTAAAGACCTTAAGCAATGGAGGTGAGACCCAGCATGTGCT 581
 QY 61 ValIleSerThrProGlnAlaSerIleValIleSerCysIleGlyGlnHisSerIleSer 80
 DB 582 ATGCTTGTACTCCTCAGGCTGCAAAAGGCAATTAAGCAAAAGACACATATGATATCA 641
 QY 81 TyrThrLeuSerArgAsnGlnThrValIleValGlyIleThrHisAspIleAspThrAsp 100
 DB 642 TATACTTGTGTCGGGCCAGACTGTGTGTGATATACATCAGACAGCAACACAGAT 701
 QY 101 MetPheGlnValIleGlySerThrGluSerProIleAspPheValIleThrAspThrIle 120
 DB 702 ATGTTTCAGATTGGCCGGTCACTAAGCCCAATGATTTTGTATGATACACAGCT 761
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 762 CCGAAGTCAAGATTAATTCATATACAGTCAGTACAAAGCATATATCAAGATTTGCC 821
 QY 141 CysArgIleValCysAspArgAsnGluProTyrThrIleAlaGlyIlePheAlaIle 160
 DB 822 TGCAGAAATCATATGGAACGGAATCTCTCTTAAGACAGCATTAATTAATGCGACAGATT 881
 QY 161 AspSerSerIleValIlePheLeuGlyGluIleValIleValIleValIleValIleValIle 180
 DB 882 GACTCATCAAAATCATCTTCTTGGGAGAGAGCTGCCAATGGAACATCAGATGGA 941
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyIlePheThr 200
 DB 942 CAGATGATGGCTTACCATATATGCTTCTTGTGATGATCCACGCAATGGGTTCCACA 1001
 QY 201 GluGluSerGlnProGlyValIleProGluIleSerValCysGlyAspValIleThrLeu 220
 DB 1002 GAAGACTCCAGCTCGAATATGAGAGAAATATCGGTGTGAGAAATGATTTAGCCCTA 1061
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyIleLeuValIleGluSerGluThrAsnValIle 240
 DB 1062 CGTGAATCCAGATCGGCTCAGCAGAGAGAGAAATGAGAAATGAAATCAATACATTA 1121
 QY 241 GlnAspGlySerLeuIleAspLeuCysGlyIleThrLeuLeuThrArgThrAlaAspGly 260
 DB 1122 CAAGATGGCTCGTTATATGACCTCTGTGTGTCACAAATGTTATGGCTGACAAAGGC 1181
 QY 261 LeuPheHisThrProThrGlnIleHisIleGlnIleAlaLeuValGlnGlnIleAsnIleAla 280
 DB 1182 CTTTCCACACTCCCTACCGTGAAGCATTTAAGCTTTTAAGACAGAGAAATCAATGACACA 1241
 QY 281 ArgProGlnCysProValIleLeuAsnThrLeuAlaPheProSerIleAsnArgGlyGlu 300
 DB 1242 CGACTCATGTGCTCCGTAGGCTTCAACACATTAAGATTTCTTAAGTATGAAGAGAAAGAC 1301
 QY 301 ValValIleGluIleValGlnProThrAlaTyrLeuSerCysGlyIleHisIleGlyIleHis 320
 DB 1302 GTTGTAGATGAAGAAACACATGAGGTATATCAATCGGACCATGATCATGCTATATCAT 1361
 QY 321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGluValGlyIleCysProMetCysArgThr 340
 DB 1362 AACTGGGAGAAACAAAGAGAAACGTGATGAGAAAGATCGTGAATGCTCTATGTGTGATGCT 1421

QY 341 ValGlyProTyrValProLeuThrPheGlyCysGluAlaGlyPheTyrValAlaAlaGly 360
 DB 1422 GTTGGTCCCTATGTTCTCTGTGGCTGTGAGCTGATTTATGTGACGCCGCGC 1481
 QY 361 ProProThiHisAlaPheThrProCysGlyValLeuValCysSerGluLysSerAlaLysTyr 380
 DB 1482 CCTCAACCCAGTCGTTTACCCGCTGTGGCATGTGTTCAGAAAGACACTGCTTAT 1541
 QY 381 TrpSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
 DB 1542 TGGTCCAGATCCCACTTCTCTATGTGATCATCTTTTCATGACGCTGCTCTTTGT 1601
 QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLeuLeuIlePheGlnGlyProIleAsp 420
 DB 1602 GCACATCAGTTGGCTGTGTGACAGGCTACATCAGACTTATTTTCAAGGACCTCTAGAC 1661
 RESULT 15
 ADF83102
 ID ADF83102 standard; cDNA; 7136 BP.
 XX ADF83102;
 AC ADF83102;
 XX 26-FEB-2004 (first entry)
 DT 26-FEB-2004 (first entry)
 XX Human Pellino homologue 1 gene, overexpressed in cancer.
 DE Human Pellino homologue 1 gene, overexpressed in cancer.
 XX Human; Pellino homologue 1; Pellino1; cancer; cytosolic; vaccine;
 KM gene therapy; gene; oncogene; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH CDS 4039..5295
 FT CDS /*tag= a
 XX MO2003100000-A2.
 XX 04-DEC-2003.
 PD 04-DEC-2003.
 XX 22-MAY-2003; 2003WO-US016049.
 PF 22-MAY-2003; 2002US-0382606P.
 XX 24-MAY-2002; 2002US-0382606P.
 PR 25-UTL-2002; 2002US-0398099P.
 XX (TUL-) TULARIK INC.
 PA L1 J, Mu D, Yang J;
 XX WPI, 2004-035118/03.
 DR P-PSDB; ADF83103.
 DR GENBANK; NM_020651.
 XX PT diagnosing a cancer in a mammal comprises determining RecQL5, CTXL,
 PT USP13, MCL1, or Pellino 1 gene copy number in a biological sample from a
 PT region of the mammal that is suspected to be precancerous or cancerous.
 PS Claim 179; SEQ ID NO 9; 174pp; English.
 XX The present sequence is that of the gene encoding human Pellino homologue
 CC 1 (Pellino1), a protein associated with the kinase domain of activated
 CC Pelle. The invention is based on the finding of the overexpression of
 CC Pellino1 and other genes (RecQL5, CTXL, USP13 and MCL1) in certain
 CC cancers, including breast cancer, colon cancer, lung cancer and ovarian
 CC cancer, and the frequent amplification of these genes in cancer cells.
 CC The genes, and their expression products, can be used diagnostically or
 CC as targets for cancer therapy. They can also be used to identify and
 CC design compounds useful in the diagnosis, prevention and therapy of
 CC tumours and cancers, in vaccine development, and in methods for
 CC determining the efficacy of a treatment regime. A claimed method for
 CC inhibiting cancer or precancerous growth, especially in ovarian or lung
 CC tissue, uses an inhibitor that interacts with Pellino1 DNA or RNA. The
 CC inhibitor is a small interfering RNA (siRNA), microRNA (miRNA), an
 CC antisense RNA, and antisense DNA, a decoy molecule, a decoy DNA, a

CC ribozyme or small molecule.
 XX Sequence 7136 BP; 1945 A; 1333 C; 1399 G; 2459 T; 0 U; 0 Other;
 SQ Alignment Scores:
 Pred. No.: 86-165 Length: 7136
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.48% Conservative: 38
 Best Local Similarity: 81.43% Mismatches: 38
 Query Match: 83.71% Indels: 2
 DB: 12 Gaps: 1
 US-10-041-030-4 (1-420) X ADF83102 (1-7136)
 QY 1 MetPheSerProGlyGlnGluGlnHisCysValaProAsnLysGluProValIleTyrGly 20
 DB 4039 ATGTTTTCCTGATCAAGAAATATAT-----CATCTTAAGACACAGTAAATATGAT 4092
 QY 21 GluLeuValIleLeuGlyTyrTrpGlnGlyAlaLeuProGlnGlyAspArgGlyTrpArgLys 40
 DB 4093 GAACCTATTGTCTTGGCTATATATGCTCTCCCAATGGCGATAGAGAAAGAGAGAA 4152
 QY 41 SerArgPheAlaLeuTyrTrpArgProLysAlaAsnGlyValIleLysProSerThrValHis 60
 DB 4153 ATTAGGTTTCTTTTGTTTTAAAGACCTAAGCAATGGGTGAGACCCAGCATGTGCAT 4212
 QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysValGlyGlnHisSerIleSer 80
 DB 4213 ATTAGGTTTCTTTTGTTTTAAAGACCTAAGCAATGGGTGAGACCCAGCATGTGCAT 4272
 QY 81 TyrThrLeuSerThrArgGlnThrValValIleGluTyrThrHisAspLysAspThrAsp 100
 DB 4273 TATACCTTGTCTCGGGCCAGACTGTGGTGTGATATATACCTAGACACACAGAT 4332
 QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValIleThrAspThrIle 120
 DB 4333 ATGTTTCAATTTGGCCGCTGACCTGAAGACCCCATGATTTTGTGATCATGACAGGTT 4392
 QY 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 4393 CCTGGAAAGTCAAGTAAATCTGATACACAGTACAGTCAAGCACTATATCAAGATTGGCC 4452
 QY 141 CysArgIleValCysAspArgGlnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
 DB 4453 TCGAGATCATATGAGAACGGAATCCTCCCTTACGACAGATTTATCTGACGAGATTT 4512
 QY 161 AspSerSerLysAsnIlePheLeuGlyGluLysAlaAlaLysTrpLysAsnProAspGly 180
 DB 4513 GACTCATCAAAAACATCTTTCTTGGGAGAAAGCTGCCAATGGAAAGACATCAGATGGA 4572
 QY 181 HisMetAspGlyLeuThrThrArgGlyValLeuValMetHisProArgGlyGlyPheThr 200
 DB 4573 CAGATGATGGCTTGACCACTAATGCTTCTTGTGATCATCCAGCAATGGGTTCCACA 4632
 QY 201 GluGlnSerGlnProGlyValIleTrpArgGluIleSerValCysGlyAspValIleThrLeu 220
 DB 4633 GAAGATCCCAAGCTCGAATATGAGAAATATCGTGTGTGGAATATTTATTTACCTTA 4692
 QY 221 ArgGluThrArgSerAlaGlnGlnAlaArgGlyLysLeuValGluSerGlnThrAsnValIle 240
 DB 4693 CCGTAAACCAAGATCGGCTACAGCAGAGAGAAATGCTGAAATGAAACCAATCAGTTA 4752
 QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrIleLeuLeuTrpArgThrAlaAspGly 260
 DB 4753 CAAGATGGCTCGTTAATTAACCTCTGTGTGCAACATTTGTATGGGCTCTCAGAAAGGC 4812
 QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnIleLeuAlaIle 280
 DB 4813 CTTTCCCAACATCTTACCGTGAAGCATTTTGAAGCTTTTAAAGCAGGAAATCAATGAGCA 4872
 QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
 DB 4873 CCACTCATGCTCCCTGTAGGCTTCAACACATAGCATTTCTTATGATGAAGAGAAAGAC 4932

```
QY 301 ValValGluGluLysGlnProTrpAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
Db 4933 GTTGTAGATGAAGAAACACCAAGGATATCTAAACTGCGCCATGATACATGACTATCAT 4992
QY 321 AenTPGIYHisArgSerAspThrGluIaAsnGluArgGluCysArgMetCysArgThr 340
Db 4993 AACTGGGGAACAAAGAAAGAGTGTGAAAAGATCGGAATGTCCTATGTGTAGTCT 5052
QY 341 ValGlyProTyrValProLeuTrpLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
Db 5053 GTTGGTCCCTATGTTCTCTGTGGCTTGATGTGAAGCTGGAATTTATGTGACGCGCGC 5112
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
Db 5113 CTTCCAAACCAATGCCGTTTAGCCCGTGTGGCAATGTGTTCAGAAAGCAAACTGCTAT 5172
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysPropheCys 400
Db 5173 TGGTCCCAAGATCCCACTTCTCATGTACTCATTACTTTTCATGCAGCCCTGCTTTTGT 5232
QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
Db 5233 GCACATCAAGTGGCTGTGTAACAGGCTACATCAGACTTATTTTCAAGACCTTAGAC 5292
```

Search completed: December 9, 2004, 13:47:55
Job time : 820 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2004, 15:53:42 ; Search time 845 Seconds
(without alignments)
2740.448 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 NPSPOBHCAPNKEPVKYG.....ATQVGEQNCIKLIFQGPID 420

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q/cg2_1/USPTO.spool.p/US10041030/runcat 09122004.095519.7785/app.query.fasta_1.583
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rmpb -MIMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -INIT=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcio -NOR=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USBR=US10041030 @CGN 1.1 912 @runcat 09122004.095519.7785
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FAPOP=6 -FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
Published Applications NA:
1: /cg2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cg2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cg2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cg2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cg2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cg2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cg2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cg2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq:*
13: /cg2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cg2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cg2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cg2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cg2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cg2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
19: /cg2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cg2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cg2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|------------------------|---------------------|
| 1 | 2290 | 100.0 | 1263 | 15 US-10-085-117-288 | Sequence 288, App |
| 2 | 2290 | 100.0 | 5579 | 14 US-10-197-666A-135 | Sequence 135, App |
| 3 | 2290 | 100.0 | 5597 | 15 US-10-085-117-287 | Sequence 287, App |
| 4 | 2290 | 100.0 | 5921 | 13 US-10-041-030-3 | Sequence 3, App |
| 5 | 2283 | 99.7 | 1263 | 9 US-09-843-905A-7 | Sequence 7, App |
| 6 | 2283 | 99.7 | 1263 | 15 US-10-317-250-7 | Sequence 7, App |
| 7 | 2283 | 99.7 | 1263 | 16 US-10-258-703-7 | Sequence 7, App |
| 8 | 2175.5 | 95.0 | 1260 | 15 US-09-843-905A-5 | Sequence 5, App |
| 9 | 2175.5 | 95.0 | 1260 | 15 US-10-317-250-5 | Sequence 5, App |
| 10 | 2175.5 | 95.0 | 1260 | 15 US-10-085-117-285 | Sequence 285, App |
| 11 | 2175.5 | 95.0 | 1260 | 16 US-10-258-703-5 | Sequence 5, App |
| 12 | 2111.5 | 92.2 | 1717 | 14 US-10-197-666A-133 | Sequence 133, App |
| 13 | 2111.5 | 92.2 | 1734 | 15 US-10-085-117-284 | Sequence 284, App |
| 14 | 1917 | 83.7 | 1257 | 9 US-09-843-905A-3 | Sequence 3, App |
| 15 | 1917 | 83.7 | 1257 | 15 US-10-317-250-3 | Sequence 3, App |
| 16 | 1917 | 83.7 | 1257 | 16 US-10-258-703-3 | Sequence 3, App |
| 17 | 1917 | 83.7 | 7136 | 13 US-10-041-030-1 | Sequence 1, App |
| 18 | 1917 | 83.7 | 7136 | 16 US-10-443-108-9 | Sequence 1, App |
| 19 | 1912 | 83.5 | 1257 | 15 US-09-843-905A-1 | Sequence 1, App |
| 20 | 1912 | 83.5 | 1257 | 15 US-10-317-250-1 | Sequence 1, App |
| 21 | 1912 | 83.5 | 1257 | 16 US-10-258-703-1 | Sequence 1, App |
| 22 | 1642 | 71.7 | 1338 | 9 US-09-843-905A-11 | Sequence 11, App |
| 23 | 1642 | 71.7 | 1338 | 15 US-10-317-250-11 | Sequence 11, App |
| 24 | 1642 | 71.7 | 1338 | 16 US-10-258-703-11 | Sequence 11, App |
| 25 | 1482.5 | 64.7 | 2508 | 15 US-10-104-047-1958 | Sequence 1958, App |
| 26 | 1064 | 46.5 | 202802 | 15 US-10-085-117-286 | Sequence 286, App |
| 27 | 1048 | 45.8 | 155579 | 15 US-10-085-117-283 | Sequence 283, App |
| 28 | 1022 | 44.6 | 541 | 15 US-10-029-386-22982 | Sequence 22982, App |
| 29 | 974 | 42.5 | 528 | 15 US-10-029-386-20504 | Sequence 20504, App |
| 30 | 808 | 35.3 | 24173 | 10 US-09-764-891-92594 | Sequence 92594, App |
| 31 | 808 | 35.3 | 24173 | 14 US-10-091-572-855 | Sequence 855, App |
| 32 | 775 | 33.8 | 465 | 9 US-09-867-550-1947 | Sequence 1947, App |
| 33 | 763 | 33.3 | 514 | 15 US-10-029-386-6790 | Sequence 6790, App |
| 34 | 763 | 33.3 | 514 | 15 US-10-029-386-9281 | Sequence 9281, App |
| 35 | 717 | 31.3 | 468 | 10 US-09-918-995-15544 | Sequence 15544, App |
| 36 | 678 | 29.6 | 377 | 9 US-09-867-550-1719 | Sequence 1719, App |
| 37 | 516 | 22.5 | 812 | 10 US-09-764-891-7046 | Sequence 7046, App |
| 38 | 516 | 22.5 | 813 | 10 US-09-764-891-7044 | Sequence 7044, App |
| 39 | 516 | 22.5 | 813 | 10 US-09-764-891-7047 | Sequence 7047, App |
| 40 | 222.5 | 9.7 | 196 | 10 US-09-764-891-9297 | Sequence 9297, App |
| 41 | 222.5 | 9.7 | 196 | 14 US-10-091-572-858 | Sequence 858, App |
| 42 | 171 | 7.5 | 5238 | 13 US-10-027-632-76172 | Sequence 76172, App |
| 43 | 171 | 7.5 | 5238 | 15 US-10-027-632-76172 | Sequence 76172, App |
| 44 | 170 | 7.4 | 1452 | 10 US-09-764-891-9295 | Sequence 9295, App |
| 45 | 170 | 7.4 | 1452 | 14 US-10-091-572-856 | Sequence 856, App |

ALIGNMENTS

RESULT 1
US-10-085-117-288
Sequence 288, Application US/10085117
Publication No. US2003023234A1
GENERAL INFORMATION:
APPLICANT: Engelhard, David
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: 529,526,0121
CURRENT APPLICATION NUMBER: US/10/085,117
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: US 09/798,586
NUMBER OF SEQ IDS: 361
SOFTWARE: Padded for Windows Version 4.0
SEQ ID NO 288
LENGTH: 1263
TYPE: DNA
ORGANISM: Homo sapiens
US-10-085-117-288

Alignment Scores:

Pred. No.: 2,556-259 Length: 1263
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 15 Gaps: 0

US-10-041-030-4 (1-420) x US-10-085-117-288 (1-1263)

```

QY 1 MetPheSerProGlyGlnGluGluHisCyba1AProAenlySGluProVallySTyrgly 20
DB 1 AAGTTTCCCTGGGCGAGAGAACACTGGCCCCCAATAAGAGCACTGTAATAACGGG 60
QY 21 GluLeuValValLeuGlyTyraAnglyAlaLeuProAenGlyAAspArgGlyArgArglys 40
DB 61 GAGCTGGTGCTGCTGGGTACATAGTGCTTTTACCCAAAGAGATAGAGACGAGAGAAA 120
QY 41 SerArgPheAlaLeuTyrysaArgProLySa1AaenGlyVallyAProSerThrValHis 60
DB 121 AGTAGATTGGCTCTACAAAGCGGCCCAAGGCAATGGTCAAAACCCAGACCGTCCAT 180
QY 61 ValIleSerThrProGlnAlaSerLySa1A1eSerCyAlyGlyGlnHisSerIleSer 80
DB 181 GTGAATATCCACGCCCGAGCATCCAAAGGCTATCAGCTGCAAGAGTCAACACAGTATATCC 240
QY 81 TyrrTrleuSerArgAaenGlnThrVallyValGluTyrrThrHisAAspLySaAProThrAp 100
DB 241 TACACTTTGTCAAGGATATAGACTGTGTGGTGAGTACACATGATTAAGATACGGAT 300
QY 101 MetPheGlnValGlyAArgSerThrGluSerProIleAAspPheValValThrAAspThrIle 120
DB 301 AAGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGCACTTGCTGTCAACAGACGANT 360
QY 121 SerGlySerGlnAaenThrAAspGluAaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 361 TCTGGCAGCGCAACACGAGCAGACGCCAGATCACACAGACACCATATCCAGTTCGCC 420
QY 141 CysAlyGlyLeValCysAAspArgAenGluProTyrrThrAlaArgIlePheAlaGlyPhe 160
DB 421 TGCAGGATCTGTGGACAGAAATGAACCTTACACAGCAGGATTTCCGCCCGGATTT 480
QY 161 AAspSerSerLySaSnIlePheLeuGlyGlyLySa1Aa1AlySTyrysaAAspProAAspGly 180
DB 481 GACTCTTCCAAAACATATTTCTTGAGAAAAGCAGCAAAAGTGGAAAACCCCGACGCG 540
QY 181 HisMetAAspGlyLeuThrThrAaenGlyVallyLeuValMetHisPProArgGlyGlyPheThr 200
DB 541 CACATGGAGTGGGCTCACTAATAAGCGCTCTGTGGTGGATGCCATCCAGAGGGGCTTTCACC 600
QY 201 GluGluSerGlnProGlyValTyrrArgGluIleSerValCyAGlyAAspValTyrrThrLeu 220
DB 601 GAGGAGTCCCAAGCCCGGGGTCTGGGCGAGATCTCTGTGTGGAGATGTGTACACCTTG 660
QY 221 ArgGluThrArgSerAlaGlnGlnAArgGlyLyLyLeuValGluSerGlnThrAaenValLeu 240
DB 661 CGAGAAACCGAGGTGGGCCAGACAGAGAAAGCTGTGGAAAAGTGAAGACCAACGTCCTG 720
QY 241 GlnAAspGlySerLeuIleAAspLeuCyArgValA1ThrLeuLeuTyrrArgThrAlaAAspGly 260
DB 721 CAGGACGGTCTCCCTCATTAAGCTGTGTGGGCCACTCTCTCTGGAAGAACAGCAGATGGG 780
QY 261 LeuPheHisThrProThrGlnLyHisIleGlnAlaLeuArgGluIleAaenAlaAla 280
DB 781 CTTTTCATCTCTCAACTCAGAAAGCATAGAAAGCCCTCCGCAAGAGATTAAACGCCGCC 840
QY 281 ArgProGlnCyAProValGlyLeuAaenThrLeuAlaAAspProSerIleAaenArgLyGlu 300
DB 841 CGGCTCAAGTGTCTGTGGGCTCAACACCTGGCTTCCAGCATCAACAGAAAGAG 900
QY 301 ValValGluGluGluGluProTrrPalA1TyrrLeuSerCyArgLyHisValHisGlyTyrrHis 320
DB 901 GTGTGGAGAGAAACAGCCCTGGGCTATATCTCAAGTGTGGCCAGGTGACGGGTACAC 960

```

```

QY 321 AenTyrrGlyHisAArgSerAAspThrGluAlaAenGluAArgGluCyAProMetCyAAspThr 340
DB 961 AACTGGGCGCATCGAGTGAACAGAGGCGCAACAGAGAGACTGTCCATGTGCGAGACT 1020
QY 341 ValGlyProTyrrValProLeuTyrrLeuGlyCyArgGluAaGlyPheTyrrValAAspAlaGly 360
DB 1021 GTGGGCCCTTATGTCTCTCTGTGGCTGTGAGGAGATTTATGTAGACCAAGGA 1080
QY 361 ProProThrHisAlaAAspThrProCyArgLyHisValCyAserGluLySerAlaArgTyrr 380
DB 1081 CCGCAACTCATGCTTTCATCTCCCTGTGACACGCTGTGAGAAAGTCTGGCAAAATAC 1140
QY 381 TrpSerGlnIleProLeuProHisGlyTyrrHisAlaPheHisAlaAlaCyAAspProPheCy 400
DB 1141 TGCTTCAGATCCCTGCTCCATGGAACCTATGCAATTTCAAGCTGTGGCTTCTGT 1200
QY 401 AlaThrGlnLeuValGlyGluGlnAaenCyAlyLeuLeuIlePheGlnGlyProIleAAsp 420
DB 1201 GCTACACAGCTGTGGTGGGAGCAAACTGCATCAATTAATTTCAGAGTCAATTCAC 1260

RESULT 2
US-10-197-666A-135
; Sequence 135, Application US/10197666A
; Publication No. US20030092037A1
; GENERAL INFORMATION:
; APPLICANT: ASAHI KASEI KABUSHIKI KAISYA
; TITLE OF INVENTION: Ekl1 phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197, 666A
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305, 884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316, 304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350, 027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 135
; LENGTH: 5579
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (177)..(1436)
US-10-197-666A-135

Alignment Scores:
Pred. No.: 2,176-258 Length: 5579
Score: 2290.00 Matches: 420
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-041-030-4 (1-420) x US-10-197-666A-135 (1-5579)
QY 1 MetPheSerProGlyGlnGluGluHisCyba1AProAenlySGluProVallySTyrgly 20
DB 177 AAGTTTCCCTGGGCGAGAGAAACACTGGCCCCCAATAAGAGCACTGTAATAACGGG 236
QY 21 GluLeuValValLeuGlyTyraAnglyAlaLeuProAenGlyAAspArgGlyArgArglys 40
DB 237 GAGCTGGTGCTGCTGGGTACATAGTGCTTTTACCCAAAGAGATAGAGACGAGAGAAA 296
QY 41 SerArgPheAlaLeuTyrysaArgProLySa1AaenGlyVallyAProSerThrValHis 60

```


Db 297 AGTAGATTGGCTCTTCAAGCGGCCCAAGCAAAATGGTGTCAACCCAGACCTTCAT 356
 Qy 61 ValIleSerThrProGlnAlaSerIleValIleSerCysValGlyGlnHisSerIleSer 80
 Db 357 GTGATATCCAGCGCCCGAGCATCCAAAGGCTATCAGCTGCAAGAGTCAACACATATATCC 416
 Qy 81 TyrThrLeuSerArgGlnGlnThrValValIleGluTyrThrHisAspIleAspThrAsp 100
 Db 417 TACACTTGTCAAGAAATAGAGTGTGGTGTGAGTACACATATATAGATAGATAGAT 476
 Qy 101 MetPheGlnValIleValSerThrGlnSerProIleAspPheValValThrAspThrIle 120
 Db 477 ATGTTTCAGGTGGGAGATCAACAGAAAGCCATATGACTTCTTGTCCAGACAGATT 536
 Qy 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 537 TCTGGCAGCAGAAACAGGACGAAAGCCAGATCACAAGAGCACCATATCCAGTTCCGC 596
 Qy 141 CysArgIleValCysAspArgGlnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
 Db 597 TCGAGATCGTGTGCAAGAGATGAACCTTACACAGCAGCGATATTCGCCCGGATTT 656
 Qy 161 AspSerSerIleAsnIlePheLeuGlyGluValAlaAlaIleTyrTrpIleAsnProAspGly 180
 Db 657 GACTCTTCCAAAACATATTTCTTGAGAAAGAGCGCAAGTGGAAAGCCCCGAGCGC 716
 Qy 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 Db 717 CACATGATGGGCTCACTAATAGCGCTGTGTGTATGATCACAAGGGGGCTTCAC 776
 Qy 201 GlnGluSerGlnProGlyValIleTrpArgGluIleSerValCysGlyAspValIleThrLeu 220
 Db 777 GAGAGATCCAGCGCGGGCTGTGGCGAGATCTGTGTGTGAGATGTACACCTTG 836
 Qy 221 ArgGluThrArgSerAlaGlnIleArgGlyValLeuValGlnSerGlnThrAsnValIle 240
 Db 837 CGAGAAACAGAGTCCGCGCCAGCAACAGAAAGCTGTGAAAGTAAAGCAACGTCCTG 896
 Qy 241 GlnAspGlySerLeuIleAspLeuCysGlyValIleThrLeuLeuTrpArgThrAlaAspGly 260
 Db 897 CAGGAGCGGCTCCCTCATGACCTGTGTGGGGCAGCTCTCTTGAGAAACAGCAGATGG 956
 Qy 261 LeuPheHisThrProThrGlnIleValIleGluAlaLeuArgGlnGluIleAsnAlaIle 280
 Db 957 CTTTTCATATCTCAACCTAGAAAGCACAATAGAAAGCCCTCGGCAAGAGATTAACGCCGC 1016
 Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgIleGlu 300
 Db 1017 CGGCTCAATGCTCTGTGGGGCTCAACACCTGCTTCCCGCATCAACAGAAAGAG 1076
 Qy 301 ValIleGlnGluValGlnProTrpAlaTyrLeuSerCysGlyValIleValHisGlyTyrHis 320
 Db 1077 GTGGTGAAGAGAAAGACCGCTGGGCATATCTCAGTTGTGGCCAGCGGATCCAG 1136
 Qy 321 AsnTrpGlyHisArgSerAspThrGlnAlaAsnGlnArgGluCysProMetCysAspGln 340
 Db 1137 AACTGGGGCATCTGAGTACACAGAGGCCCAAGAGAGAGTGTCCATGTGCAGAGCT 1196
 Qy 341 ValIleProTyrValProLeuTrpLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
 Db 1197 GTGGGCGCTATGTGCTCTGTGGCTGTGAGGAGAGATTTTATATGACGAGAG 1256
 Qy 361 ProProThrHisAlaPheThrProCysGlyValIleValCysSerGlnIleSerAlaIle 380
 Db 1257 CCGCCAACTCATCTTCACTCCCTGTGACAGTGTGGCTCGGAGAAAGTCTGCAAAATAC 1316
 Qy 381 TrpSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaIleCysProPheCys 400
 Db 1317 TGGTTCAGATCCCGTGTGCTCAAGAACTCAAGCATTTCAACCTTGTCCCTTCTGT 1376
 Qy 401 AlaThrGlnLeuValGlyGlnAsnCysValIleValLeuIlePheGlnGlyProIleAsp 420

Db 1377 GCTACACAGCTGTGGGAGCAAAAGTCAATTAATTTTCCAGGTCCAAATTGAC 1436
 RESULT 3
 US-10-085-117-287
 ; Sequence 287, Application US/10085117
 ; Publication No. US2003023234A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
 ; FILE REFERENCE: 529452000121
 ; CURRENT APPLICATION NUMBER: US/10/085,117
 ; PRIOR FILING DATE: 2002-02-27
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; NUMBER OF SEQ ID NOS: 361
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 287
 ; LENGTH: 5597
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-085-117-287
 Alignment Scores:
 Pred. No.: 2,18e-258 Length: 5597
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 15 Gaps: 0
 US-10-041-030-4 (1-420) x US-10-085-117-287 (1-5597)
 Qy 1 MetPheSerProGlnGlnGluIleValCysAlaProAsnIleGluProValIleTyrGly 20
 Db 177 ATGTTTCCCGCGCGCAGAGGAAACCTGCGCCCAATTAAGAGCAGTGAATACGGG 236
 Qy 21 GlnLeuValIleGluTyrThrAsnGlyAlaLeuProAsnGlyAspArgGlyTyrGly 40
 Db 237 GAGCTGGTGTCTCGGTACATGTGCTTAAACCAATGAGATAGAGAGCGAGAA 296
 Qy 41 SerArgPheAlaLeuTyrTyrAspProIleValAsnGlyValIleProSerThrValHis 60
 Db 297 AGTAGATTGGCTCTTCAAGCGGCCCAAGCAAAATGGTGTCAACCCAGACCTTCAT 356
 Qy 61 ValIleSerThrProGlnAlaSerIleValIleSerCysValGlyGlnHisSerIleSer 80
 Db 357 GTGATATCCAGCGCCCGAGCATCCAAAGGCTATCAGCTGCAAGAGTCAACACATATATCC 416
 Qy 81 TyrThrLeuSerArgGlnGlnThrValValIleGluTyrThrHisAspIleAspThrAsp 100
 Db 417 TACACTTGTCAAGAAATAGAGTGTGGTGTGAGTACACATATATAGATAGATAGAT 476
 Qy 101 MetPheGlnValIleValSerThrGlnSerProIleAspPheValValThrAspThrIle 120
 Db 477 ATGTTTCAGGTGGGAGATCAACAGAAAGCCATATGACTTCTTGTCCAGACAGATT 536
 Qy 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 537 TCTGGCAGCAGAAACAGGACGAAAGCCAGATCACAAGAGCACCATATCCAGTTCCGC 596
 Qy 141 CysArgIleValCysAspArgGlnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
 Db 597 TCGAGATCGTGTGCAAGAGATGAACCTTACACAGCAGCGATATTCGCCCGGATTT 656
 Qy 161 AspSerSerIleAsnIlePheLeuGlyGluValAlaAlaIleTyrTrpIleAsnProAspGly 180
 Db 657 GACTCTTCCAAAACATATTTCTTGAGAAAGAGCGCAAGTGGAAAGCCCCGAGCGC 716
 Qy 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 Db 717 CACATGATGGGCTCACTAATAGCGCTGTGTGTATGATCACAAGGGGGCTTCAC 776

```

QY 201 GIUGLUSERG1PROG1YVALTRPARGLUILESERVALCYSG1YASPVATYTRThleu 220
DB 777 GAGGAGTCCCAAGCCCGGGGTCTGGCGGAGATCTCTGTCTGTGGAGATGTGACACTTG 836
QY 221 ARGGLUTHRARGSERVALG1NG1NARG1LYLYLEUVALG1USERG1UTHRANVALLEU 240
DB 837 CGAGAAACAGGTCGGCCGACGAAACGAGGAAAGCTGTGGAAAGTGAACCAACGCTCG 896
QY 241 G1NAPRG1YSERLEU1LEAPLEUCYSG1YATHTHEULEUTRPARGTHRLAASPV 260
DB 897 CAGGAGGCTCCCTCACTTACCTGTGTGGGCGCACTCTCTCTGGAGAACACAGATGGG 956
QY 261 LEUPHEH1STHRPROTHG1NLYSH1LEGLUALALEUARG1NG1UILEASNA1A1A 280
DB 957 CTTTTTCATCTCCAACTCAGAAAGCACATAGAAGCCCTCCGGCAGAGATTAACCGCCGC 1016
QY 281 ARGPROG1NCYARPROVALG1LYLEUANTHRLEUALAPHEPROSER1LEASNA1RG1YSG1U 300
DB 1017 CGGCGCTCAGTGTCTGTGGGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 1076
QY 301 VAL1VALG1UG1ULYRG1NPROTPRALATYRLEUSERCYSG1YH1SVALH1SG1LYTRH1S 320
DB 1077 GTGGTGAGAGAGAGAGAGCCCTGGGCATCTCATGTTGTGGCCACGTGCAAGGTAACAC 1136
QY 321 AANTPRG1YH1ARGSERAPRTHRG1UALAANG1UARG1UCYARPROKCYASARGTHR 340
DB 1137 AACTGGGGCCATCGAGATGACACGAGGCGCAACGAGAGAGATGCTCCATGTGCGAGAT 1196
QY 341 VAL1GYPROTYRVALPROLEUTRLEUG1LYCYSG1UALG1LYPHETRYVALASPV 360
DB 1197 GTGGGCGCCCTATGTGCTCTCTGTGGCTGTGGAGGAGATTTATGTACACGACAGA 1256
QY 361 PROPROTHH1ALAPHERTHRPROCYSG1YH1SVALCYASERG1ULYUSER1ALYSTR 380
DB 1257 CCGCACTCACTGCTTCTCTCTGTGGAGACAGTGTGCTCGAGAGAGCTGCAAAATAC 1316
QY 381 TRPSERGIN1LEPROLEUPROH1SG1YTHRH1SALAPHEH1SALALA1CYARPROPHCY 400
DB 1317 TGGTCTCAGATCCCGTGTGCTCATGSACTCATGATTTCAAGCTGCTTGCCTTTCTGT 1376
QY 401 ALATHRG1NLEUVALG1LYUG1NANCYG1LYLYLEU1LEPHG1NG1YPRO1LEAP 420
DB 1377 GCTACACAGCTGGTGGGAGAGAACTGATCAATTAATTTCCAAAGTCAATTGAC 1436

RESULT 4
US-10-041-030-3
; Sequence 3, Application US/10041030
; Publication No. US20020150934A1
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Mu, David
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian
; FILE OF INVENTION: Pellino Polypeptides and Polynucleotides
; FILE REFERENCE: 018781-006810US
; CURRENT APPLICATION NUMBER: US/10/041,030
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/259,502
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5921
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (501)..(1763)
; OTHER INFORMATION: human pellino 2
US-10-041-030-3

```

Alignment Scores:

| Pred. No.: | 2,376-258 | Length: | 5921 |
|------------------------|-----------|---------------|------|
| Score: | 2290.00 | Matches: | 420 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 13 | Gaps: | 0 |

US-10-041-030-4 (1-420) x US-10-041-030-3 (1-5921)

```

QY 1 MetPheSerProG1YGL1UG1N1H1SCYSA1APROANLYRG1UPROVAL1YSTRG1Y 20
DB 501 ATGTTTCCCTCGGCGCAGAGAGAACATCGCCGCCCAATAAGAGACGATGAATACCGG 560
QY 21 G1LEUVAL1LEUG1LYTRPARG1VALLEUPROANG1YASPVATYRARGH1Y 40
DB 561 GAGCTGGTGTCGCTGGGTACATGTCCTTACCAATGGAGATAGAGACGAGGAGAA 620
QY 41 SERATPHEALALEUTRYLYEARPROLYSALAAANG1YVALLYSPROSERTHRYALH1S 60
DB 621 AGTAAATTTGCCCTTACAGCGGCGCCCAAGCAATGCTCAACCAAGACACGCTCAT 680
QY 61 VAL1LESERTHRPROG1NALASERYAL1LESERYLYSG1YGLNH1SER1LESER 80
DB 681 GTGATATCCAGCGCCCAAGGATCCAAAGGCTTACAGCTCAAGGTCACACAGTATATCC 740
QY 81 TYRTHRLEUSERARGANG1NTHRYVAL1VALG1UTYRTHH1ASPVLYASPVTHRAP 100
DB 741 TACACTTTGTCAAGAAATCAAGCTGTGGTGGAGTACACATGATPAGATACGAT 800
QY 101 METPHEG1NVALG1YARGSERTHRG1USERPRO1LEAPHEVAL1YTHRAPTHR1LE 120
DB 801 ATGTTTCAAGTGGCGAGATCAACAGAAAGCCCTTACGCTTGTGTACAGACAGATT 860
QY 121 SERG1YSERGINASERTHRAPG1UALAG1N1LETHRG1N1SERTHR1LESERARGPHEAL 140
DB 861 TCTGGCACACGAAACACGAGAGAGCCAGATCAACAGACACACCATATCCAGGTTCCGC 920
QY 141 CYSAAG11EVALCYASPARARGANG1UPROTYRTHRALAARG1LEPHAL1ALAG1Y 160
DB 921 TGCAGAGATGCTGTGGAGAGAGATACCTTACAGACAGATTTGCGCGCGGATTT 980
QY 161 ASPSER1LYASNA1LEPHELEUG1YGL1ULYSA1A1ALYSTR1LYASNP1ROASPV 180
DB 981 GACTCTTCCAAATAATATTTCTTGGAGAAAGGACGAAAGTGAAACCCCGACGCG 1040
QY 181 HISWETAPRG1YLEUTHRTHRAPNG1YVALLEUVAL1METH1SPROARG1YGLYTRH 200
DB 1041 CACATGAGATGGGCTCATTAATGCGTCTGTGGATGACATCCACGAGGGGCTTTCAC 1100
QY 201 GIUGLUSERG1NPROG1YVALTRPARGLUILESERVALCYSG1YASPVATYTRThleu 220
DB 1101 GAGGAGTCCCAAGCCCGGGGTCTGGCGGAGATCTCTGTCTGTGGAGATGTGACACTTG 1160
QY 221 ARGGLUTHRARGSERVALG1NG1NARG1LYLYLEUVALG1USERG1UTHRANVALLEU 240
DB 1161 CGAGAAACAGGTCGGCCGACGAAACGAGGAAAGCTGTGGAAAGTGAACCAACGCTCG 1220
QY 241 G1NAPRG1YSERLEU1LEAPLEUCYSG1YATHTHEULEUTRPARGTHRLAASPV 260
DB 1221 CAGGAGGCTCCCTCATTAATGAGCTGTGTGGGCGCACTCTCTCTGGAGAACAGCAATGG 1280
QY 261 LEUPHEH1STHRPROTHG1NLYSH1LEGLUALALEUARG1NG1UILEASNA1A1A 280
DB 1281 CTTTTTCATCTCCAACTCAGAAAGCACATAGAAGCCCTCCGGCAGAGATTAACCGCCGC 1340
QY 281 ARGPROG1NCYARPROVALG1LYLEUANTHRLEUALAPHEPROSER1LEASNA1RG1YSG1U 300
DB 1341 CGGCGCTCAGTGTCTGTGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 1400
QY 301 VAL1VALG1UG1ULYRG1NPROTPRALATYRLEUSERCYSG1YH1SVALH1SG1LYTRH1S 320
DB 1401 GTGGTGAGAGAGAGAGAGCCCTGGGCATATCTCATGTTGTGGCCAGTGTGACAGGTAACAC 1460

```

QY 321 AERTTPGLYHIAERSEARPHTRGUAIAAENGUAIGUCYSPROMETCYBATGTHR 340
 DB 1461 AACTGGGGCGCATCGGATGACACGAGGCGCAAGAGGAGTGTCCATGTGCAGAGACT 1520
 QY 341 VALIGLPRGYTVALIProleuTPrleuGLYCYSGUAAAGLYPHEGYTVALAAspAGLY 360
 DB 1521 GTGGGCGCCCTATGTGCTCTCTGCTTGGCTGTGAGCGAGATTATATGACGAGGA 1580
 QY 361 PROPTORHIAIAIAPHERHPRCYSGLYHIAVALCYSESGILUYSESEALALYTYR 380
 DB 1581 CCGCCAACTCATGCTTCACTCCCTGTGACAGTGTGCTCGGAGAAAGTGCAGAAATAC 1640
 QY 381 TRPSERGINLEPRoleuProhiSGLYTHRIHIAIAPHEHIAIAAACYSPROPHECY 400
 DB 1641 TGGTTCAGATCCCGTTGCTCATGGAACATGACATTCACCTGCTGCTTCTGT 1700
 QY 401 ALATHGLILEUVALIGLYUGLIHIAENCYVILIEVLEUULEPHEGINGIProileAP 420
 DB 1701 GCTACACAGCTGTGGGAGCAAACTGCATCAAAATTAATTTCCAAAGGTCCAAATTGAC 1760
 RESULT 5
 US-09-843-905A-7
 ; Sequence 7, Application US/09843905A
 ; Patent No. US20020168683A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Cosman, David J.
 ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 ; FILE REFERENCE: 2990-A
 ; CURRENT APPLICATION NUMBER: US/09/843,905A
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,198
 ; PRIOR FILING DATE: 2000-04-28
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 1263
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-843-905A-7
 Alignment Scores:
 Pred. No.: 1,7e-258 Length: 1263
 Score: 2283.00 Matches: 419
 Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 1
 Query Match: 99.69% Indels: 0
 DB: 9 Gaps: 0
 US-10-041-030-4 (1-420) x US-09-843-905A-7 (1-1263)
 QY 1 MetPhSeSerProGlynglugluHIASeCyAIAProAsnLySLeuProValIlyrTyrgLy 20
 DB 1 AAGTTTTCCTCCGCGCAAGAGAACTGCGCCCAATPAGAGCAAGGAAATACGG 60
 QY 21 GIULIeUVALIeUGLYTYrAENGlyAIAleuProAsnLyAAspArgLyA-gar-gLy 40
 DB 61 GAGCTGTGTGCTCGGAGTACATGTGTCTTACCAATGAGATAGAGAGCGAGAGAA 120
 QY 41 SeArPheAlaIeUlyrTyrsArProlYsAlaAsnGLYValIlyrProSeArThrValIly 60
 DB 121 AGTAGATTGGCCCTTCAAGCGGCCCAAGCAAAATGTGTCAAAACCCAGCACCGTCCAT 180
 QY 61 ValIleSeArThrProGlnAlaSeLySAlaIleSeCyAlySGLYGlnHIASeRileSer 80
 DB 181 GTGATATTCACGCGCCAGCATCCAGGCTATCATGCTGCAGAAAGTCAACATATATTC 240
 QY 81 TyrThLeuSeArAgaEngInThrValIValIgluTyrThRIAspLySArPThrAP 100
 DB 241 TACACTTTGTCAAGGATCATGACTGTGTGTGAGATACACATATTAAGATACGAGAT 300
 QY 101 MetPheGlnValIglYrSeArThrGlySeRProIleAspPheValIlyrAspThrIle 120

DB 301 ATGTTTCAGTGGGAGATCAACAGAAAGCCCTATGCACTGTTTTCACAGACAGATT 360
 QY 121 SerGlySeRGIAsnThrAspGluAIAglnIleThrGlnSeRThRIleSeArPheAla 140
 DB 361 TCTGGCAGCAGAAACCGAGCGAAGCCAGATCAACAGAGCAACATATTCAGGTTCCGC 420
 QY 141 CyAAGIleValIcyAAspArgAEngIlyProTyrThrAlaArgIlePheAlaIagLyPhe 160
 DB 421 TCCAGGATGTGTGCACAGGATGAACTTACACAGCACAGATATTCGCCCGGATTT 480
 QY 161 AsPSeSerLyAsnIlePheLeuGLYGLYsAlaIAlaLyETrPlyAsnProAspGly 180
 DB 481 GACTTTCAGAAACATATTTCTGGAGTAAAGGCGCAAAAGTGGAAACCCGACGGC 540
 QY 181 HIAmeLAspGlyLeuThrThrAsnGLYValIleUValmeChIAProArgLyglYpHeThr 200
 DB 541 CACATGGATGGGCTCATCTATAGGCGTCTGTGTATGCATCCACGAGGGGCTTCACC 600
 QY 201 GIULISeRGIProGlyValITPrArgIuIleSeValIcySGLYAspValIlyrThrLeu 220
 DB 601 GAGAGTCCACGCCGGGCTGTGGCGAGATCTGTCTGTGAGATGTACACCTTG 660
 QY 221 ArGGLuThrArGSeArAIAglnGlnArGGLYsLeUValIGlySeRGIuThrAsnValI 240
 DB 661 CAGAAACAGAGTGGCCAGCAACAGAAAGCTGTGAAAGTGAAGCAACATCTCTG 720
 QY 241 GlnAspGlySeRleuIleAspLeuCYSGlyAIArThrLeuLeuTrPArGThrIAAspGly 260
 DB 721 CAGGACGGCTCCCTATTAACCTGTGTGGGCACTCTCTCGAGAAACAGAGATGGG 780
 QY 261 LeuPheHIArThrProThrGlnLySHIleGluAlaLeuArgGlnIuIleAsnAlaIA 280
 DB 781 CTTTTTCATATCCCACTCAGAAAGCATAGAAAGCCCTCCGACAGAGATTAAACGCCGC 840
 QY 281 ArGProGlnIcyProValIGlyLeuAsnThrLeuAlaPheProSeRThrAsnArgLySGL 300
 DB 841 CGGCTCATGTGCTGTGGGAGCTCAACACCTGCTTCCAGCATCAACAGAAAGAG 960
 QY 301 ValIAlIGluGluLySLeuProTrPArLyTyLeuSeRcySGLYHIAValHISGLYTHRI 320
 DB 901 GTGGTGAAGAGAGAGAGCCCTGGGCAATATTCAGTTGTGGCAGTCGACGGGTACAC 960
 QY 321 AERTTPGLYHIAERSEARPHTRGUAIAAENGUAIGUCYSPROMETCYBATGTHR 340
 DB 961 AACTGGGGCGCATCGGATGACACGAGGCGCAAGAGAGAGTGTCCATGTGCAGAGACT 1020
 QY 341 VALIGLPRGYTVALIProleuTPrleuGLYCYSGUAAAGLYPHEGYTVALAAspAGLY 360
 DB 1021 GTGGGCGCCCTATGTGCTCTCTGCTTGGCTGTGAGCGAGATTATATGACGAGGA 1080
 QY 361 PROPTORHIAIAIAPHERHPRCYSGLYHIAVALCYSESGILUYSESEALALYTYR 380
 DB 1081 CCGCCAACTCATGCTTCACTCCCTGTGACAGTGTGCTCGGAGAAAGTGCAGAAATAC 1140
 QY 381 TRPSERGINLEPRoleuProhiSGLYTHRIHIAIAPHEHIAIAAACYSPROPHECY 400
 DB 1141 TGGTTCAGATCCCGTTGCTCATGGAACATGACATTCACGCTGCTGCTTCTGT 1200
 QY 401 ALATHGLILEUVALIGLYUGLIHIAENCYVILIEVLEUULEPHEGINGIProileAP 420
 DB 1201 GCTACACAGCTGTGGGAGCAAACTGCATCAAAATTAATTTCCAAAGGTCCAAATTGAC 1260
 RESULT 6
 US-10-317-250-7
 ; Sequence 7, Application US/10317250
 ; Publication No. US20030165945A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Cosman, David J.
 ; APPLICANT: Li, Xiaoxia
 ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 ; FILE REFERENCE: 2990-B

QY 221 ArgGluThrArgSerAlaGlnGlnAlaArgGlyLysLeuValGlyIleuSerGluThrAsnValLeu 240
Db 658 CGAGNACACAGGTCGGCCCGACGAGAGGGAAAGCTGTGGAAAGTGAAGCAACACGTCCTG 717
QY 241 GlnAspGlySerLeuIleAspLeuCySGlyValAlaThrLeuLeuTrpArgThrAlaAspGly 260
Db 718 CAAGACGGGCTCCCTCATTTGACCTGTGTGGGGCCACTCTCTCGAGAAACCGCAGATGGC 777
QY 261 LeuPheHisThrProThrGlnLysHisGlyIleGluAlaLeuArgGlnGlnIleLeuAlaAla 280
Db 778 CTTTTCACGCTCTTCTCAAGAGCACTATGAAGCCCTCCGACGAGATCAATGCAAGCC 837
QY 281 ArgProGlnCyAspProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
Db 838 CGACCCCAAGTGCCCGTGGGCTTTAACCCCTGGCTTCCCAAGATCAACCGGAGAA 897
QY 301 ValValGluGluLysGlnProTrpAlaTyrLeuSerCySGlyHisValHisGlyTyrHis 320
Db 898 GTGGTGAAGAGAGAGAGCCCTGGGCAATGAGCTGGCGCATGTGCACGGCTACAC 957
QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCyAspProMetCyAspTrp 340
Db 958 AGCTGGGGCCATCGAGAGCAAGCGGAAAGCCACAGAGAGAGTGTCCATGTGCAGACT 1017
QY 341 ValGlyProTyrValProLeuTrpLeuGlyCySGlyValGlyPheTyrValAspAlaGly 360
Db 1018 GTGGGCCCCCTACGCTCTCTGTGGCTGTGGCTGTGAAGATTTTATGTGATGGGA 1077
QY 361 ProProThrHisAlaPheThrProCySGlyHisValCySerGluLysSerAlaLysTyr 380
Db 1078 CCCCCAACCTCAGCTTTCACCCCTCGGGCAGCTGTGTTCAGAAAGCTCGCAAGTAC 1137
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCyAspProPheCyS 400
Db 1138 TGGTCGCAATCCCATCTGCCCGACGAGAGCGACGCCATTACGCCCTGTGCTCTGC 1197
QY 401 AlaThrGlnLeuValGlyGlnAsnCySileTyrLeuIlePheGlnGlyProIleAsp 420
Db 1198 GCCACGACCTGGTGTGTGAACGAAGCTGATCAATTGATTTTCCAAAGTCCAGTGAC 1257

RESULT 9
US-10-317-250-5
; Sequence 5, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; APPLICANT: Li, Xiaoxia
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-B
; CURRENT APPLICATION NUMBER: US/10/317,250
; CURRENT FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-317-250-5

Alignment Scores:
Pred. No.: 7, 84e-246 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.384 Conservative: 10
Best Local Similarity: 95.008 Mismatches: 10
Query Match: 95.008 Indels: 1
DB: 15 Gaps: 1

US-10-041-030-4 (1-420) x US-10-317-250-5 (1-1260)

QY 1 MetPheSerProGlyGlnGlnGluHisCySAlaProAsnLysGluProValLysTyrGly 20
Db 1 ATGTTTTCCTCCGCGCCAGAGAAACCAAGCCGCCCAACAGAGAGCCGGTGAATACGGG 60

QY 21 GlnLeuValValLeuGlyTyrAsnGlyValAlaLeuProAsnGlyAspArgGlyValArgLys 40
Db 61 GAGCTGGTGTCTCTGGGATACATGTGCTTTACTTAATGTGACAGGGGCGAGAGAA 120
QY 41 SerArgPheAlaLeuTyrIleAspArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db 121 AGCAGATTGGCTCTATPAGCGGACCTAAGCGAGTGGGTCAAAACCGACCAATCCAC 180
QY 61 ValIleSerThrProGlnAlaSerLysValIleSerCySlyGlyGlnHisSerIleSer 80
Db 181 ATGGTCTCACACCAAGCGCTCCAGAGCCATCACTCCAGAGGACATCAGCATATCG 240
QY 81 TyrThrLeuSerArgAsnGlnThrValValValGlyTyrThrHisAspLysAspThrAsp 100
Db 241 TACAGCTTGTACGAGAGCCAGCGGTAGTGGAGTACACACAGATTAACACAGCAGC 300
QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
Db 301 ATGTTTCAGGTGGCAGGTCAACAGAAAGCCCATTTGACTTGTGTGCACAGACAGGTT 360
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 361 TCCGGCGGTCAAGAAAGAAAT--GCCAGATCACACAGACCATCTTAAGTTTCGCA 417
QY 141 CysArgIleValCyAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 418 TCCAGATCTGTGTGACAGAGACGACCATACACACGATTTTCCGGCAGAGATTG 477
QY 161 AspSerSerLysAsnIlePheLeuGlyGlyLysAlaAlaLysTyrLysAsnProAspGly 180
Db 478 GATTCCTCCAAAATAATCTTCTTGTGAGAGAAAGCAGCAAAATGAGAAACCTGATGCA 537
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProAspGlyGlyPheThr 200
Db 538 CACATGATGAGACTCACTACCAATGTGTCTTAAGATGACCCCGCAAGAGAGCTTCACC 597
QY 201 GlnGluSerGlnProGlyValTrpArgGluIleSerValCySGlyAspValTyrThrLeu 220
Db 598 GAGGAATCCAGCTCGAGTCTGAGAGAGATCTGTCTGTGGGAGATGTGTACACTTG 657
QY 221 ArgGluThrArgSerAlaGlnGlnAlaArgGlyLysLeuValGlyIleuSerGluThrAsnValLeu 240
Db 658 CGAGAGACCAAGGTGCGGCCAGCAGAGAGGAAAGCTGTGAGAAAGTGAAGCAACCTCCG 717
QY 241 GlnAspGlySerLeuIleAspLeuCySGlyValAlaThrLeuLeuTrpArgThrAlaAspGly 260
Db 718 CAAGACGGGCTCCCTCATTTGACCTGTGTGGGGCCACTCTCTCGAGAAACCGCAGATGGC 777
QY 261 LeuPheHisThrProThrGlnLysHisGlyIleGluAlaLeuArgGlnGlnIleLeuAlaAla 280
Db 778 CTTTTCACGCTCTTCTCAAGAGCACTATGAAGCCCTCCGACGAGATCAATGCAAGCC 837
QY 281 ArgProGlnCyAspProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
Db 838 CGACCCCAAGTGCCCGTGGGCTTTAACCCCTGGCTTCCCAAGATCAACCGGAGAA 897
QY 301 ValValGluGluLysGlnProTrpAlaTyrLeuSerCySGlyHisValHisGlyTyrHis 320
Db 898 GTGGTGAAGAGAGAGAGCCCTGGGCAATGAGCTGGCGCATGTGCACGGCTACAC 957
QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCyAspProMetCyAspTrp 340
Db 958 AGCTGGGGCCATCGAGAGCAAGCGGAAAGCCACAGAGAGAGTGTCCATGTGCAGACT 1017
QY 341 ValGlyProTyrValProLeuTrpLeuGlyCySGlyValGlyPheTyrValAspAlaGly 360
Db 1018 GTGGGCCCCCTACGCTCTCTGTGGCTGTGGCTGTGAAGATTTTATGTGATGGGA 1077
QY 361 ProProThrHisAlaPheThrProCySGlyHisValCySerGluLysSerAlaLysTyr 380
Db 1078 CCCCCAACCTCAGCTTTCACCCCTCGGGCAGCTGTGTTCAGAAAGCTCGCAAGTAC 1137

```

QY 381 TrpSerGlnIleProIeuProHisGlyThrHisAlaPheHisAlaIaCysProPheCys 400
DB 1138 TGGTCCAGATCCCACTGCCCCACGAGACGACGCGTTTCATCCCGCTGCTCGCTCGC 1197
QY 401 AlaThrGlnLeuValGlyGlnGlnAnCysIleLeuLeuIlePheGlnGlyProIleAsp 420
DB 1198 GCCACGACGCTGTTGTGTGAACAGAACTGCATCAATTAATTTTCCAGGTCCAGTGCAC 1257

RESULT 10
US-10-085-117-285
; Sequence 285, Application US/10085117
; Publication No. US200302334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 285
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-085-117-285

Alignment Scores:
Pred. No.: 7,846-246 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.388 Conservative: 10
Best Local Similarity: 95.008 Mismatches: 1
Query Match: 15 Gaps: 1

US-10-041-030-4 (1-420) x US-10-085-117-285 (1-1260)
QY 1 MetPheSerProGlyGlnGlnGlnHisCysAlaProAsnIleGluProValIleGlyTyr 20
DB 1 ATGTTTCCCGGCGGACGAGGAGAACCCAGCGCCCAACAGAGCGCGTGAATACCGG 60
QY 21 GluLeuValIleGlnGlyTyrAsnGlyAlaLeuProAsnIleAspArgGlyTyrGly 40
DB 61 GAGCTGCTGCTCTGGGCTACATATGCTTACTTAATGCTGACAGGCGGACGAGGAA 120
QY 41 SerArgPheAlaLeuTyrIleAspArgProIleAsnGlyValIleProSerThrValHis 60
DB 121 AGCAGATTGCTCTTATTAAGCGGACCTACGCCAGTGTGTCAAAACCCAGCAATCCAC 180
QY 61 ValIleSerThrProGlnAlaSerIleValIleSerCysLeuGlyGlnHisSerIleSer 80
DB 181 ATGCTTCACACCAACAGCGGTCTCAAGCCATCAGTCCAGAGACATCAGCATATCG 240
QY 81 TyrThrLeuSerArgAsnGlnThrValIleValIleGlyTyrThrHisAspIleAspThrAsp 100
DB 241 TACACCTTGTCAACGAGCGACGAGTGTGTGAGTATACACACATTAAGACACAGGAC 300
QY 101 MetPheGlnValIleGlyArgSerThrGlnSerProIleAspPheValIleThrAspThrIle 120
DB 301 ATGTTTCAGGTGGCGGCTCAACAGAAAGCCCATTAAGTCTTGCTGCTACAGACGAGTT 360
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 361 TCCGGCGGTCAAGAACAGAT--GCCAGATCACACAGAGCAACATCTCTAGTTCCGA 417
QY 141 CysArgIleValIleCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
DB 418 TCGAGGATGCTGTGACAGGAGCAACGACATATACAGCAGCATATTCCGGGCGAGATTC 477
QY 161 AspSerSerIleAsnIlePheLeuGlyGlnValAlaIleTyrThrIleAspArgGly 180

```

```

DB 478 GATTTCCAAAATATCTTTCTGAGAGAAAGCAAAATGAAAAACCTGATGCA 537
QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleValIleMetHisProArgGlyGlyPheThr 200
DB 538 CACATGATGATGACTCACTACCAATGAGTGTCTTAGTATGACACCCGCAAGAGGCTTCACC 597
QY 201 GluIleuSerGlnProGlyIleTyrAspArgIleIleSerValCysGlyIleAspValIleThrIleu 220
DB 598 GAGGAATCCACGCTGAGTGTGAGAGAGATCTGTGTGTGGGATGTATACACTTG 657
QY 221 ArgGluThrArgSerAlaGlnGlnIleArgGlyIleLeuValIleSerGluThrAsnValIleu 240
DB 658 CAGAGACCAAGTCCGCGCCACGAGAGGAGAAAGCTGTGTGAAGTGAACCAACGCTCTG 717
QY 241 GlnAspGlySerIleuIleAspLeuCysGlyAlaIleThrLeuLeuTyrArgThrAlaAspGly 260
DB 718 CAAGACGCTCCCTCATTAACCTGTGTGGGCGCACTCTCTCGAGAACCCGAGATGAC 777
QY 261 LeuPheHisThrProThrGlnIleHisIleGluAlaLeuArgGlnIleIleAsnAlaIa 280
DB 778 CTTTTCACGCTCTTACTGAGAGCAGATGAAGCCCTCCGACGAGATCAATGACGCC 837
QY 281 ArgProGlnCysProValIleLeuAsnThrLeuAlaPheProSerIleAsnArgIleGlu 300
DB 838 CACCCCAATGCCCCGTGGGCTTAAACACCTTGCTTCCACGATCAACCGAGAGAA 897
QY 301 ValIleGlnGlnIleProIleProTyrIleLeuSerCysGlyIleValIleGlyTyrHis 320
DB 898 GTGTGTGAAGAGAGAGAGCGCTGGGCTATCTGAGCTGGGCGCATGTGACGCGCTACAC 957
QY 321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysAspThr 340
DB 958 ACTGTGGGCGATCGAGCGCACCGGAGCAACGAGAGGAGAGTGTCCATGTGCAGAGCT 1017
QY 341 ValGlyProTyrValProIleuTyrLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
DB 1018 GTGGGCTTACGCTCTCTGTGGCTGGGCTGTGAGCGAGATTTATGTATGCTGCGGA 1077
QY 361 ProProThrHisAlaPheThrProCysGlyIleValCysSerGluIleSerAlaIleTyr 380
DB 1078 CCCCCAATCAGCTTTCACCCCTGGGCGGACGCTGTTCAGAAAGTCTGCAAGTAC 1137
QY 381 TrpSerGlnIleProIeuProHisGlyThrHisAlaPheHisAlaIaCysProPheCys 400
DB 1138 TGGTCCAGATCCCACTGCCCCACGAGACGACGCGTTTCATCCCGCTGCTCGCTCGC 1197
QY 401 AlaThrGlnLeuValGlyGlnGlnAnCysIleLeuLeuIlePheGlnGlyProIleAsp 420
DB 1198 GCCACGACGCTGTTGTGTGAACAGAACTGCATCAATTAATTTTCCAGGTCCAGTGCAC 1257

RESULT 11
US-10-258-703-5
; Sequence 5, Application US/10258703
; Publication No. US20040034199A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-US
; CURRENT APPLICATION NUMBER: US/10/258,703
; PRIOR FILING DATE: 2002-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-258-703-5

Alignment Scores:
Pred. No.: 7,846-246 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.388 Conservative: 10

```


Best Local Similarity: 95.00% Mismatches: 10
 Query Match: 95.00% Indels: 1
 DB: 16 Gaps: 1

US-10-041-030-4 (1-420) x US-10-258-703-5 (1-1260)

```

QY      1 MetPheSerProGlyGlnGluGluHisCyValAProAsnLysGluProValLysTyrGly  20
DB      1 ATGTTTTCCTCCGGGCGCAGAGAGAACCCAGCGCCCAACAAGAGCGGTGAAATACGGG  60
QY      21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyValArgGly  40
DB      61 GAGCTGTGTCTCTGGGGGTACATGTGCTTAACTTAATGTGTACACGGGGCGAGGAAA  120
QY      41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis  60
DB      121 AGCAGATTGCCCCCTATAGCGGACCTAGCCAGTGTGTCAAAACCCAGCAATCCAC  180
QY      61 ValLysSerThrProGlnAlaSerLysAlaLysSerCyValLysGlyGlnHisSerLys  80
DB      181 ATGCTCTCACACACAGCGCGCTCAAGGCATACGTCCAGAGAGCATCACAGCATATCG  240
QY      81 TyrThrLysSerArgAsnGlnThrValValValGluTyrThrHisAspLysAspThrAsp  100
DB      241 TACACGTTGTCCACGAGCCAGACCGGTAGTGTGTGAGTACACACCATTAAGACACGGAC  300
QY      101 MetPheGlnValGlyArgSerThrGluSerProLysPheValValThrAspThrIle  120
DB      301 ATGTTTCAGGTGGCGAGGTCAACAGAAAGCCCATTTGACTGTGTGTCAACAGACGGTT  360
QY      121 SerGlySerGlnAsnThrAspGluValGlnIleThrGlnSerThrLysSerArgPheAla  140
DB      361 TCCGGCGGTGACAGAACGAAAT---GCCAGATACACAGAGCAACCATCTCTAGGTTCCGA  417
QY      141 CyValGlyLeuValCyAspArgAsnGluProTyrThrAlaArgLysPheAlaArgLys  160
DB      418 TCCAGAGTGTGTGTGACAGAGACGAGCATATACAGCAGCATATTCCTCCGACAGATTC  477
QY      161 AspSerSerLysAsnLysPheLeuGlyGluValAlaLysThrLysAsnProAspGly  180
DB      478 GATTCTTCCAAAATAATCTTTCTTGGAGAAAGACGAAATATGGAAAACCTGTATGGA  537
QY      181 HisMetAspGlyLeuThrThrAsnGlyValLysValMetHisProArgGlyGlyPheThr  200
DB      538 CACATGATGAGACTACATCAATGATGTCCTATGATGATCAACCCGAAAGAGGCTTCAAC  597
QY      201 GluGluSerGlnProGlyValTyrArgGluLysSerValCyGlyAspValTyrThrLeu  220
DB      598 GAGGAATCCACGCTGAGTCTGAGAGAGATCTGTCTGTGTGGGATGTGTACACCTTG  657
QY      221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu  240
DB      658 CGAGAGACCAAGTCCGCCACAGACAGAGGAGAGCTGTGTGAAAAGTGAACCAACGTTCTG  717
QY      241 GlnAspGlySerLeuLysAspLeuCyGlyAlaThrLeuLeuTyrArgThrAlaAspGly  260
DB      718 CAAGACGGCTCTCTATGACCTGTGTGGGCCACTCTCTGTGAGAAACCCGAGATGCG  777
QY      261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluLysAsnAla  280
DB      778 CTTTTTCAAGCTCTTACTCAGAAAGCAGATAGAGCCCTCCGACAGAGATCAATGACGCG  837
QY      281 ArgProGlnCyProValGlyLeuAsnThrLeuAlaPheProSerLysAsnArgLysGlu  300
DB      838 CGAACCCCAATGCCCCCTGGGCTTAAACACCTGGCTTCCCAAGCATCAACCGAAAGGA  897
QY      301 ValValGluGluLysGlnProTyrAlaTyrLysSerCyGlyHisValHisGlyTyrHis  320
DB      898 GTGGTGAAGAGAGAGACGCTGGGCAATCTGAGCTGGGCCATGTGACAGGATCAACAC  957
QY      321 AsnTyrGlyHisLysSerAspThrGluAlaAsnGluArgGluCyProMetCyAspArgThr  340
DB      958 ACCTGGGCGCATCGAGCGACCGGAAAGCAAGAGAGGAGTGTCTCCATGTGACGAGACT  1017

```

```

QY      341 ValGlyProTyrValProLeuThrLeuGlyCyGlyAlaArgLysPheTyrValAspAlaGly  360
DB      1018 GTGGGCCCCCTACGTCCTCTCTGAGCTGTGTGAGGACGAGATTTATGTGATCGGGGA  1077
QY      361 ProProThrHisAlaPheThrProCyGlyHisValCySerGluLysSerAlaLysTyr  380
DB      1078 CCCCCAATCTCACGCTTTCACCCCTCGGGGACGCTGTGTTCAGAAAGTGTGCCAAGTAC  1137
QY      381 TrpSerGlnLysProLeuProHisGlyThrHisAlaPheHisAlaAlaCyAspPheCys  400
DB      1138 TGTGTGAGATCCCACTGCCCCACGAAACGACGCGTTTCATGCGCTGTCTCGTTCGC  1197
QY      401 AlaThrGlnLeuValGlyGluGlnAsnCyHisLysLeuLysPheGlnGlyProLysAsp  420
DB      1198 GCCACGACGCTGTGTGTGAAACGAACTGCATCAATGATTGTTCCAAAGTCCATGTGAC  1257

RESULT 12
US-10-197-666A-133
; Sequence 133, Application US/1019766A
; Publication No. US20030092037A1
; GENERAL INFORMATION:
; APPLICANT: ASAHI KASEI KABUSIKI KAISYA
; TITLE OF INVENTION: Elkl phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 133
; LENGTH: 1717
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)..(1474)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1288..1301
; OTHER INFORMATION: n-A or C or G or T
US-10-197-666A-133

Alignment Scores:
Pred. No.: 4,25e-238 Length: 1717
Score: 2111.50 Matches: 388
Percent Similarity: 95.24% Conservative: 12
Best Local Similarity: 92.36% Mismatches: 19
Query Match: 92.21% Indels: 1
DB: 14 Gaps: 1

US-10-041-030-4 (1-420) x US-10-197-666A-133 (1-1717)
QY      1 MetPheSerProGlyGlnGluGluHisCyValAProAsnLysGluProValLysTyrGly  20
DB      1 ATGTTTTCCTCCGGGCGCAGAGAGAACCCAGCGCCCAACAAGAGCGGTGAAATACAGG  277
QY      21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyValArgGly  40
DB      278 GAGCTGTGTCTCTGGGGGTACATGTGCTTAACTTAATGTGTACACGGGGCGAGGAAA  337
QY      41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis  60

```



```

Db      338 AGCAGATTACCTCTATTAAGCGACTACGCCAGTGGGTCAAAACCCAGCAATATCCAC 397
      61 ValIleSerThrProGlnAlaSerIleValIleSerCysValGlyGlnHisSerIleSer 80
      398 ATGGTCTCCACACACAGAGCGGTCCAGGCGCATCAGCTCCAGAGCATCATCAGCATATGCG 457
      81 TyrThrLeuSerArgAsnGlnThrValValValGluTyrThrHisAspIleAspThrAsp 100
      458 TACACGCTGTCCAGGAGCAGAGAGGTAGTGGAGTATACACACATTAAGCACAGAC 517
      101 MetPheGlnValAlaYarSerThrGlnSerProIleAspPheValValThrAspThrIle 120
      518 ATGTTTCAGGTGGGCGAGGTCAACAGAAAGCCCATTTGACTCTGTGTCACAGACGGTT 577
      121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
      578 TCCGGCGGTCAAGAACAGAT--GCCAGATCACACAGACCATCTTCTAGTTCCGA 634
      141 CysArgIleValCysAspArgAsnGlnProTyrThrAlaArgIlePheAlaIleGlyPhe 160
      635 TCGAGGATCGTGTGACAGAGAACGACCATACAGACGCGCATATTCGGCGCAGGATTTC 694
      161 AspSerSerIleAsnIlePheLeuGlyGluValAlaIleTyrTrpIleAsnProAspGly 180
      695 GATTCTTCCAAAATAATCTTTCTTGGAGAGAAAGCAGCAAAATGGAAAACCTGTATGGA 754
      181 HisMetAspGlyLeuThrThrAsnGlyValIleValMetHisAspArgGlyGlyPheThr 200
      755 CACATGATGAGACTCATCAACAGGTGCTGTGTATGACCCGAGAGAGGCTTCACCC 814
      201 GluGluSerGlnProGlyValIleTyrAspGlnIleSerValCysGlyAspValIleThrLeu 220
      815 GAGGAATCCCAAGCTGAGTCTGAGAGAGATCTCTGTCTGTGGGATGATACACCTTG 874
      221 ArgGluThrArgSerAlaGlnIleAspGlyValLeuValGluSerGlnThrAsnValIleu 240
      875 CAGAGACCAAGTCCGCCAGCAGAGAGGAAAGCTGTGAGAAAGTACCAACGCTCCG 934
      241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTrpArgThrAlaAspGly 260
      935 CAAGAGCGGTCCCTCATTTGACCTGTGGGGCCACTCTCTCGAGAAACCGCAGATGGG 994
      261 LeuPheHisThrProThrGlnIleValIleGlnAlaLeuArgGlnIleAsnAlaIle 280
      995 CTTTTCACGCTCTCTACTCAGAAAGCACAATAGAAAGCCCTCCGCGAGAGATCAATGACGCC 1054
      281 ArgProGlnCysProAlaGlyLeuAsnThrLeuAlaPheProSerIleAsnArgIleGlu 300
      1055 CAGCCCAAGTCCCGCTGGGCTTAAACCTGAGCTCTCCCGACATCAACCGGAAAGAA 1114
      301 ValIleGluGluValProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
      1115 GTGGTGAAGAGAGAGAGCGCTGGGCAATCCTGAGCTGGGCAATGACCGCTACAC 1174
      321 AsnTrpGlyHisArgSerAspThrGlnAlaAsnGlnArgGluCysProMetCysArgThr 340
      1175 CACTGGGGGCATCGAGCAGACAGGAAAGCCAGAGAGGAGTGTCCCATGTGCATGACT 1234
      341 ValIleProTyrValProLeuTrpLeuGlyCysGlyAlaGlyPheTyrValAspAlaGly 360
      1235 GTGGGCGCTACCTCTCTGCTGGCTGTGAGGACAGATTTATGTCCGANNNNNN 1294
      361 ProProThrHisIleAspThrProCysGlyHisValCysSerGlnIleSerAlaValTyr 380
      1295 NNNNNNNCATGTTTTCACCCCTCGGGCAGTGTCTTCAAGAGCTGCGCAAGTAC 1354
      381 TrpSerGlnIleProLeuProHisGlyTyrHisIlePheHisIleAlaCysProPheCys 400
      1355 TGGTCCCAATCCCACTGCCCGCAGAGACGCGCTTTCAACCCCGCTGTCCGTTCTGC 1414
      401 AlaThrGluLeuValGlyGluGlnAsnCysIleLeuLeuIlePheGlnGlyProIleAsp 420

```

```

Db      1415 GCCAGCGACTGTGTGTGAACAGACTGCATCAATTATTTTCCAGGTCCAGTGCAGC 1474
RESULT 13
US-10-085-117-284
; Sequence 284, Application US/10085117
; Publication No. US2003023234A1
; GENERAL INFORMATION:
; APPLICANT: Moritz, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: Variation
; LOCATION: (1)...(1734)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-284

```

Alignment Scores:

```

Pred. No.: 4,31e-238 Length: 1734
Score: 2111.50 Matches: 388
Percent Similarity: 95.24% Conservative: 12
Best Local Similarity: 92.38% Mismatches: 19
Query Match: 92.21% Indels: 1
DB: Gaps: 1

```

US-10-041-030-4 (1-420) x US-10-085-117-284 (1-1734)

```

QY      1 MetPheSerProGlyGlnGluGlnHisCysAlaProAsnIleGluProValIleValTyrGly 20
      218 ATGTTTCCCGCGGCGCAGAGAGAACCCAGCGCCCAACAGAGCGGTGAATACAG 277
      21 GluLeuValIleGlyTyrTrpAsnGlyAlaLeuProAsnGlyAspArgIleArgAlaGly 40
      278 GAGCTGTGCTCTGGGATGACATGCTTAACTTAATGTGACAGGGGCAAGAGAA 337
      41 SerArgPheAlaLeuTyrIleAspProIleAsnGlyValIleProSerThrValHis 60
      338 AGCAGATTACCTCTATTAAGCGACTACGCCAGTGGTGCMAACCCAGCAATATCCAC 397
      61 ValIleSerThrProGlnAlaSerIleValIleSerCysValGlyGlnHisSerIleSer 80
      398 ATGGTCTCCACACACAGAGCGGTCCAGGCGCATCAGCTCCAGAGCATCATCAGCATATGCG 457
      81 TyrThrLeuSerArgAsnGlnThrValValValGluTyrThrHisAspIleAspThrAsp 100
      458 TACACGCTGTCCAGGAGCAGAGAGGTAGTGGAGTATACACACATTAAGCACAGAC 517
      101 MetPheGlnValAlaYarSerThrGlnSerProIleAspPheValValThrAspThrIle 120
      518 ATGTTTCAGGTGGGCGAGGTCAACAGAAAGCCCATTTGACTCTGTGTCACAGACGGTT 577
      121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
      578 TCCGGCGGTCAAGAACAGAT--GCCAGATCACACAGACCATCTTCTAGTTCCGA 634
      141 CysArgIleValCysAspArgAsnGlnProTyrThrAlaArgIlePheAlaIleGlyPhe 160
      635 TCGAGGATCGTGTGACAGAGAACGACCATACAGACGCGCATATTCGGCGCAGGATTTC 694
      161 AspSerSerIleAsnIlePheLeuGlyGluValAlaIleTyrTrpIleAsnProAspGly 180
      695 GATTCTTCCAAAATAATCTTTCTTGGAGAGAAAGCAGCAAAATGGAAAACCTGTATGGA 754

```

```

QY 181 HisMeAspGlyLeuThrThraEnglyValLeuValMeHisProarglyGlyPheThr 200
DB 755 CACATGGAGTCACTCACTCAACGGTCTCGTGTATGACCCGCAAGAGCTTCACC 814
QY 201 GluGlySerGlnProGlyValItrpArgGluIleSerValCysGlyAspValTyrThrIleu 220
DB 815 GAGGATATCCACCTGGAGTCTGAGAGAGATCTCTGTGTGGGATGTATACCTTG 874
QY 221 ArgGlyThrArgSerAlaGlnItrpArgGlyValLeuValGlySerGlnThrArgValIleu 240
DB 875 CGAGAGACCAAGGTGGCCGACGAGAGAGGAGGAGTGTGTGAAGTGAAGCAACGCTTG 934
QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTrrpArgThrAlaAspGly 260
DB 935 CAAGACGGCTCCCTCATTTACCTGTGTGGGCACTCTCTGTGAGGAACCCGAGATGGC 994
QY 261 LeuPheHisThrProThrGlnIlyshisIleGlyAlaLeuArgGlnIleAsnAlaAla 280
DB 995 CTTTTCACGCTCTACTACAGAGCACATAGAAAGCCCTCCGACAGAGATCAATGACGC 1054
QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArglySerGlu 300
DB 1055 CCAACCCCAAGTCCCGTGGGCTCTTACACCTGGCTTCCCAAGCATCAACCGAAGGA 1114
QY 301 ValValGluGlyLeuGlnProTrrpAlaTyrLeuSerCysGlyValIshisGlyTyrHis 320
DB 1115 GTGGTGAAGAGAAAGAGAGCCCTGGGATACCTGAGCTGGGCACTGTGACCGCTACCC 1174
QY 321 AsnTrrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
DB 1175 CACTGGGGGCATCGGAGCAGACGAGAACCAAGAGAGAGTGTCCCATGTGCATGACT 1234
QY 341 ValGlyProTrrpValProLeuTrrpLeuGlyCysGluAlaGlyPheTyrAlaAspAlaGly 360
DB 1235 GTGGGCCCCCTAGCTCTCTGTGGCTGTGAGGCGAGTTTATGTCAANNNNNN 1294
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlnIlySerAlaIlyTyr 380
DB 1295 NNNNNNNCTCATGTTTTCACCCCTGGGCGAGCTGTGTGAGAGAGCTGCCAAGTAC 1354
QY 381 TrrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
DB 1355 TGGTGGCAGATCCCACTGGCCCAAGGAGCAGCGCTTTCAGCGCCCTGTCCGTTCTGC 1414
QY 401 AlaThrGlnLeuValGlyGlnAsnCysIleIlyLeuIlePheGlnIlyProIleAsp 420
DB 1415 GCCAGCAGCTGGTGTGTAACAGAACTCAATTAATTAATTTCCAAAGTCCAGTGCAC 1474

RESULT 14
US-09-843-905A-3
; Sequence 3. Application US/09843905A
; Patent No. US2002016863A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-843-905A-3

Alignment Scores:
Pred. No.: 2,236-215 Length: 1257
Score: 1917.00 Matches: 342
Percent Similarity: 90.48% Conservative: 38

```

```

Best Local Similarity: 81.43% Mismatches: 38
Query Match: 83.71% Indels: 2
DB: 9 Gaps: 1

US-10-041-030-4 (1-420) x US-09-843-905A-3 (1-1257)
QY 1 MetPheSerProGlyGlnGluIleHisCysValaProLeuValIlyPheTrrpGly 20
DB 1 ATGTTTCTCCGATCAAGAAATATCAT-----CCATGTAAAGCAACGATAAATATGTT 54
QY 21 GluLeuValIleValGlyTrrpEnglyAlaLeuProAsnGlyAspArglyValArgArgly 40
DB 55 GAACCTATGCTTGTGATATATGATCTCTCCAAACGGGATAGAGAGGAGGAAA 114
QY 41 SerArgPheAlaLeuTrrpValArgProValAlaEnglyValIlyPheProSerThrPheHis 60
DB 115 AGTAGGTTTGTCTTGTAAAGACTTAAGGCAATAGGAGGAGAACCCGACCTGTGCAT 174
QY 61 ValIleSerThrProGlnAlaSerIlyValaIleSerCysIlyGlyGlnHisSerIleSer 80
DB 175 ATTGCTTGATCTCCCAAGCTGCAAGAGCATTAAGCAACAAAGACACATATGATATCA 234
QY 81 TrrpThrLeuSerArgAsnGlnThrValValValGluTrrpHisAspIlyAspThrAsp 100
DB 235 TATACTTATATCTCGGACCAGACTGTGTGTGATATATCAATGACAGACACCGAT 294
QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIleThrAspThrIle 120
DB 295 ATGTTTCAGATTGGCGGTTCACATGAAGGCCCATATTTGTGATGATGACACGGTT 354
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 355 CCGTGAAGTCAAGAAATTCATGATACACAGTCAGACAGAACAGACATATCAAGATTGCC 414
QY 141 CysArgGlyIleValCysAspArgAsnGluProTrrpAlaArgIlePheAlaAlaGlyPhe 160
DB 415 TGCAGAAATCAATATGAAACGGAATCTCTTTCACGACGAGATTTAATGTCGACAGGTTT 474
QY 161 AspSerIlyAsnIlePheLeuGlyGlyIlyValaIlyIleTrrpIlyAsnProAspGly 180
DB 475 GACTATCAAAATAATCTTTCTTGGAGAGAGGCTGCCAATGGAAAGATCAAGATGGA 534
QY 181 HisMeAspGlyLeuThrThraEnglyValLeuValMeHisProarglyGlyPheThr 200
DB 535 CAGATGGATGGTTGACACATATATGCTTCTGTGATGATCAACGCAAGGGTTCCACA 594
QY 201 GluGlySerGlnProGlyValItrpArgGluIleSerValCysGlyAspValTyrThrIleu 220
DB 595 GAAGACTCCAAAGCCCTGGAATATGAGAGAAATATCGTGTGGGAAATGATATTGACCTTA 654
QY 221 ArgGlyThrArgSerAlaGlnIleArgIlyIlyLeuValGlySerGlnThrAsnValIleu 240
DB 655 CGTGAACCAAGATCGGCTCAGAGAGAGAGAAATGTGTGAATTTGAATCCATATAGTTTA 714
QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTrrpArgThrAlaAspGly 260
DB 715 CAAGATGCTGTGTTAATGACCTGTGTGCAACATTTGATGGCGATGCAAGAGGC 774
QY 261 LeuPheHisThrProThrGlnIlyshisIleGlyAlaLeuArgGlnIleAsnAlaAla 280
DB 775 CTTTCCACACTCCCTACCGGTGAAGCATTTAGAGCTTTAAGACAGGAATCATCATACACA 834
QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArglySerGlu 300
DB 835 CCACTCAGTGCCTGTAGGGTTCAACACACTAGCATTTCTCAGATATAGAGGAGAAAGAC 894
QY 301 ValValGluGlyLeuGlnProTrrpAlaTyrLeuSerCysGlyValIshisGlyTyrHis 320
DB 895 GTTGTAGATGAAGAAACAAACATGGGTATATCTAAACATCGGCAATGTACATGCTATCAT 954
QY 321 AsnTrrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
DB 955 AACTGGGGAACCAAGAGAGAGAGTGTGATGATGCTTATGTGTGTGCT 1014

```

```

QY 341 ValGlyProTyrValProLeuTrrPLeuGlyCyseGluAlaGlyPheTyrValAspAlaGly 360
Db 1015 GTTGGGCCCTATATCTCTGTGGCTTGTGATGTAAGCTGATTTATGTGGACGCCGCC 1074
QY 361 ProProthrhniAlaPheThrProCyseGlyVhiValCyseSerGluysSerAlaLysTyr 380
Db 1075 CTTCCAAACCAATCGCTTTACCCCGTGTGGCAATGTGTTCAGAAAGAACAACTGCCTAT 1134
QY 381 TrpSerGlnIleProLeuProhiAglyThrniAlaPhehiAlaAlaCyseProPheCyse 400
Db 1135 TGGTCCCAAGTCCCACTTCTCATGTGATCATACTTTTCATGACGCTGTCTCTTTTGT 1194
QY 401 AlaThrGlnLeuValGlyGluGlnAsnCyseIleLysLeuIlePheGlnGlyProIleAsp 420
Db 1195 GCACATCAGTTGGCTGTGTAACAAGCTATCATCAGACTTATTTTTCAGAGCACTCTAGAC 1254

RESULT 15
US-10-317-250-3
; Sequence 3, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; APPLICANT: Li, Xiaoxia
; TITLE OF INVENTION: HUMAN PBLINO POLYPEPTIDES
; FILE REFERENCE: 2990-B
; CURRENT APPLICATION NUMBER: US/10/317,250
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-317-250-3

Alignment Scores:
Pred. No.: 2,23e-215 Length: 1257
Score: 1917.00 Matches: 342
Percent Similarity: 90.48% Conservative: 38
Best Local Similarity: 81.43% Mismatches: 38
Query Match: 83.71% Indels: 2
DB: 15 Gaps: 1

US-10-041-030-4 (1-420) x US-10-317-250-3 (1-1257)
QY 1 MetPheSerProGlyGlnGluGlnIleCyseAlaProAsnIleGluProValLysTyrGly 20
Db 1 ATGTTTTCCTCTGATCAAGAAATCAT-----CAATCTAAAGCACCAAGTAATAATATG 54
QY 21 GluLeuValIleGluGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGlyLys 40
Db 55 GAACATATGTCTTAGATATATATGATCTCTCCAAACCGTATAGAGGAAGAGAGAA 114
QY 41 SerArgPheAlaLeuTyrIleValArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db 115 AGTAGCTTGTCTTTGTTTAAAGACCTAAGGCAAAAGCGGTGAAGCCACGACTGTGCAT 174
QY 61 ValIleSerThrProGlnIleSerLysAlaIleSerCyseLysGlyGlnIleSerIleSer 80
Db 175 ATTGCTTGATCTCTCAAGGCTGCAAGGCAATAGCAACAAACCAACCACTATGATCA 234
QY 81 TyrThrLeuSerArgAsnGlnThrValIleValGlyTyrThrniAspLysAspThrAsp 100
Db 235 TATACTTTATCTCGGGCCAGACTGTGTGTGTAATATATCTCATGACAGCAACACGAT 294
QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIleThrAspThrIle 120
Db 295 ATGTTCACAGTTGGCCGCTGACGTAAGCCCAATGATTTTGTAGTAACGACACGCTT 354
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140

```

```

Db 355 CTTGAAATGCAAGTAATTTGTATACAGTCAAGTACAAAGCACTATTCAGATTTGCC 414
QY 141 CyseArgIleValCyseAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 415 TGCAGAAATCATATGTGAACGGAAATCTCCCTTTACAGCAACGGATTTATGCTGCAGGGTT 474
QY 161 AspSerSerLysAsnIlePheLeuGlyGlyValAlaLysTrrLysAsnProAspGly 180
Db 475 GACTCATCAAAAACATCTTTCTTGGGGAAGAGCGCCAAATGGAAGACATCAGATGGA 534
QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleuValMetHisProArgGlyGlyPheThr 200
Db 535 CAGATGATGGCTTGACCACTAATGTGTTCTTGATGATCATCCAGCAATGGGTTTCACA 594
QY 201 GluGluSerGlnProGlyValTrrArgGluIleSerValCyseGlyAspValTyrThrLeu 220
Db 595 GAAGACTCAAGCCTGGAATATGAGAAATAATGCGTGTGGGAATGATTTAGCCTA 654
QY 221 ArgGluThrArgSerAlaGlnIleArgGlyLysLeuValGluSerGluThrAsnValLeu 240
Db 655 CGTGAAACCAAGATCGGCTCAGCAGAGAGAAATAATGCGAATTTGAAACCAATCAGTTA 714
QY 241 GlnAspGlySerLeuIleAspLeuCyseGlyValIleThrLeuThrArgThrAlaAspGly 260
Db 715 CAAATGGCTCGTTAATTTAATCTCTGTGGTCAACATGTTATGGCTTCTGCAGAAAGGC 774
QY 261 LeuPhehiSerProThrGlnIleLysVhiIleGluAlaLeuArgGlnGluIleAsnAlaAla 280
Db 775 CTTTCCCAACATCCCTACCGTGAAGCACTTTAAGAGCTTTAAGACAGAAATCAATGCAGCA 834
QY 281 ArgProGlnCyseProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
Db 835 CGACCTCAGTGCCTGTAGGTTTCAACACATTCATTTCTTATGTGAAGAGAAAGAC 894
QY 301 ValIleGluGlyLysGlnProTrrAlaTyrLysSerCyseGlyVhiSerValIleGlyTyrHis 320
Db 895 GTTGTAGATGAAAACAAACCAATGGATATCTTAACCTGGCGCATGTACATGGCTATCAT 954
QY 321 AsnTrrGlyVhiArgSerAspThrGlnIleAsnGluArgGlyCyseProMetCyseArgThr 340
Db 955 AACTGGGGAAGAAACAAAGAAAGACGTGATGGCAAAAGATCGTAATGTCTTATGTAGTCT 1014
QY 341 ValGlyProTyrValProLeuTrrPLeuGlyCyseGluAlaGlyPheTyrValAspAlaGly 360
Db 1015 GTTGGTCCCTATATCTCTGTGGCTTGTGATGTGAAGCTGATTTATGTGACGCCGCC 1074
QY 361 ProProthrhniAlaPheThrProCyseGlyVhiValCyseSerGluysSerAlaLysTyr 380
Db 1075 CTTCCAAACCAATCGCTTTACCCCGTGTGGCAATGTGTTCAGAAAGAACAACTGCCTAT 1134
QY 381 TrpSerGlnIleProLeuProhiAglyThrniAlaPhehiAlaAlaCyseProPheCyse 400
Db 1135 TGGTCCCAAGTCCCACTTCTCATGTGATCATACTTTTCATGACGCTGTCTCTTTTGT 1194
QY 401 AlaThrGlnLeuValGlyGluGlnAsnCyseIleLysLeuIlePheGlnGlyProIleAsp 420
Db 1195 GCACATCAGTTGGCTGTGTAACAAGCTATCATCAGACTTATTTTTCAGAGCACTCTAGAC 1254

```

Search completed: December 10, 2004, 00:00:36
 Job time : 872 secs

This Page Blank (uspto)

CL

P.D. 03-09-2001
1-2 (2)

ID AAH34231 standard; cDNA; 864 BP.

XX
AC AAH34231; XP-002235785
XX

DT 03-SEP-2001 (first entry)

XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:1313.

XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.

XX
OS Homo sapiens.

XX
PN WO200122920-A2.

XX
PD 05-APR-2001.

XX
PP 28-SEP-2000; 2000WO-US26524.

XX
PR 29-SEP-1999; 99US-0157137.

XX
PR 03-NOV-1999; 99US-0163280.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;

XX
DR WPI; 2001-235357/24.

XX
DR P-PSDB; AAG74826.

XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers .

XX
PS Claim 1; Page 3049; 9803pp; English.

XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.

CC
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX
SQ Sequence 864 BP; 247 A; 142 C; 157 G; 314 T; 3 other;

| | |
|--|-----|
| ctgcttaatt gaagtgtaat ataggtttta gaagtgttac ctgcagttct atggttttct | 60 |
| ttcacttctt tcttttttta aagccattct gttcttttga tgtgcttgaa aggggtgtgtg | 120 |
| attacaccat tgttaatgct gggtaaaaaa taactttctc cagccttgcc tcataacagt | 180 |
| ggaattttctg atagacaaac cacaggaact tgaatttaag ccaaatccat ctccatccct | 240 |
| ttactgtcaa tcttctgtcc cagtgtttta gcccttttgc cttaggttat gatgcgcctc | 300 |
| cttctgtgcg accaatgaga cgaattcaga acatttttaa aataatctaa gcatcattga | 360 |
| agcagtaaca caaaaaaaag gttcatttat tcttttttga tataacttac atcctttcaa | 420 |
| ataagtcttt gccctcatga agaatccat caggaagata aggaataata gtattttcca | 480 |
| gttttgcttg acagttttctt aacaaaccaa aataacata atgaaggaa agatgtttct | 540 |

This Page Blank (uspto)